

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 5, 2005, 15:08:33 ; Search time 694 Seconds

(without alignments)
3184.733 Million cell updates/sec

Title: US-10-077-111-13

Perfect score: 2047
Sequence: 1 MVKLHTLADHGDVNCACF.....LTPNRLKMAINRWLETHQK 384

Scoring table:

BLISSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 433806 segs, 2877871033 residues

Total number of hits satisfying chosen parameters: 3483234

Minimum DB seg length: 0
Maximum DB seg length: 200

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=x1h
-Q=/cg2_1/USPTO.spool/US10077111/runat_04022005_121142_7236/app_query.fasta_1.583
-DB=Published Applications NA -QFMT=fastlap -SUFFIX=p2nszlm200.rnpb
-MINMATCH=0.1 -LOOPEXT=0 -UNITs=bits -START=1 -END=1
-MATRIX=bloms62 -TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct
-THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=200
-USER=US10077111_QCGN_1_1_628 @runat_04022005_121142_7236 -NCPU=6 -ICPU=3
-NO MMP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELCK=100 -LONGLOG
-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA: *
1: /cg2_6/ptodata/2/pubpna/US07_PUBCOMB.seq: *
2: /cg2_6/ptodata/2/pubpna/PCT NEW PUB.seq: *
3: /cg2_6/ptodata/2/pubpna/US06 NEW PUB.seq: *
4: /cg2_6/ptodata/2/pubpna/US06_PUBCOMB.seq: *
5: /cg2_6/ptodata/2/pubpna/US07_NEW PUB.seq: *
6: /cg2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq: *
7: /cg2_6/ptodata/2/pubpna/US08 NEW PUB.seq: *
8: /cg2_6/ptodata/2/pubpna/US08_PUBCOMB.seq: *
9: /cg2_6/ptodata/2/pubpna/US09_PUBCOMB.seq: *
10: /cg2_6/ptodata/2/pubpna/US09_PUBCOMB.seq: *
11: /cg2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq: *
12: /cg2_6/ptodata/2/pubpna/US09_NEW PUB.seq: *
13: /cg2_6/ptodata/2/pubpna/US10_PUBCOMB.seq: *
14: /cg2_6/ptodata/2/pubpna/US10_PUBCOMB.seq: *
15: /cg2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq: *
16: /cg2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq: *
17: /cg2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq: *
18: /cg2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq: *
19: /cg2_6/ptodata/2/pubpna/US10_NEW PUB.seq: *
20: /cg2_6/ptodata/2/pubpna/US11 NEW PUB.seq: *
21: /cg2_6/ptodata/2/pubpna/US60 NEW PUB.seq: *
22: /cg2_6/ptodata/2/pubpna/US60_PUBCOMB.seq: *

Pred. No. is the number of results predicted by the chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	300	14.7	180	9	US-09-864-761-27687	Sequence 27687, A
C 2	210	10.3	132	9	US-09-864-761-30542	Sequence 30542, A
C 3	150	7.3	92	15	US-10-106-699-3381	Sequence 3381, Ap
C 4	93.5	4.6	174	16	US-10-029-386-26571	Sequence 26571, A
C 5	87	4.3	150	16	US-10-029-386-18677	Sequence 18677, A
C 6	82	4.0	173	9	US-09-864-761-29419	Sequence 29419, A
C 7	79.5	3.9	159	17	US-10-305-720-417	Sequence 417, App
C 8	78.5	3.8	155	16	US-10-029-386-16111	Sequence 16111, A
C 9	72.5	3.5	197	9	US-09-998-598-205	Sequence 205, App
C 10	72.5	3.5	197	9	US-09-998-598-2327	Sequence 2327, App
C 11	72.5	3.5	197	18	US-10-425-115-24663	Sequence 24663, A
C 12	71.5	3.5	169	17	US-10-424-599-124934	Sequence 124934, A
C 13	67.5	3.3	131	16	US-10-029-386-25501	Sequence 25501, A
C 14	65.5	3.2	167	14	US-10-060-036-1810	Sequence 1810, Ap
C 15	65	3.2	131	9	US-09-864-761-23956	Sequence 23956, A
C 16	64	3.1	118	16	US-10-029-386-16384	Sequence 16384, A
C 17	64	3.1	142	16	US-10-029-386-16305	Sequence 16305, A
C 18	63.5	3.1	167	17	US-10-424-599-37936	Sequence 37936, A
C 19	62.5	3.1	156	10	US-09-814-353-5015	Sequence 5015, Ap
C 20	62.5	3.1	156	10	US-09-814-353-11307	Sequence 11307, A
C 21	62	3.0	36	13	US-10-077-111-19	Sequence 19, App1
C 22	61	3.0	123	18	US-10-425-115-136942	Sequence 136942, A
C 23	61	3.0	181	9	US-09-920-100A-1407	Sequence 1407, Ap
C 24	61	3.0	181	13	US-10-033-528-1407	Sequence 1407, Ap
C 25	61	3.0	181	16	US-10-099-926-1407	Sequence 1407, Ap
C 26	61	3.0	189	16	US-10-029-386-15506	Sequence 15506, A
C 27	61	3.0	196	17	US-10-424-599-42401	Sequence 42401, A
C 28	60.5	3.0	177	9	US-09-867-701-6419	Sequence 6419, Ap
C 29	60	2.9	189	9	US-09-738-826-3303	Sequence 3303, Ap
C 30	59.5	2.9	103	9	US-09-864-761-21318	Sequence 21318, A
C 31	59.5	2.9	189	11	US-09-922-293-2560	Sequence 2560, Ap
C 32	59	2.9	103	9	US-09-969-373-651	Sequence 651, App
C 33	59	2.9	115	11	US-09-987-899-6165	Sequence 6165, Ap
C 34	59	2.9	177	16	US-10-029-386-21065	Sequence 21065, Ap
C 35	59	2.9	183	16	US-10-029-386-21433	Sequence 21433, A
C 36	59	2.9	133	18	US-10-425-115-39652	Sequence 39652, A
C 37	58.5	2.9	118	17	US-10-377-803-15	Sequence 15, App1
C 38	58.5	2.9	180	18	US-10-437-963-28604	Sequence 28604, A
C 39	58	2.8	165	18	US-10-437-963-58869	Sequence 58869, A
C 40	58	2.8	192	18	US-10-437-963-90128	Sequence 90128, A
C 41	57.5	2.8	145	17	US-10-242-535A-54429	Sequence 54429, A
C 42	57.5	2.8	145	17	US-10-085-783A-54429	Sequence 54429, A
C 43	57	2.8	141	17	US-10-282-122A-5293	Sequence 5293, Ap
C 44	57	2.8	141	17	US-10-282-122A-5300	Sequence 5300, Ap
C 45	57	2.8	151	9	US-09-864-761-30033	Sequence 30033, A

ALIGNMENTS

RESULT 1
US-09-864-761-27687/C
; Sequence 27687, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Weisheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aemica-X-1
; CURRENT APPLICATION NUMBER: US/09/864, 761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366

```

; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 27687
; LENGTH: 180
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006501.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.88
; OTHER INFORMATION: EXPRESSED IN PETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.81
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.73
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.72
; OTHER INFORMATION: EST HUMAN HIT: AL040518.1, EVALUATE 1.00e-94
; OTHER INFORMATION: NT HIT: U43139.1, EVALUATE 1.10e-01
; OTHER INFORMATION: SWISSPROT HIT: Q16760, EVALUATE 6.80e-02
US-09-864-761-27687

Alignment Scores:
Pred. No.: 2,16e-26 Length: 180
Score: 300.00 Matches: 58
Percent Similarity: 98.31% Conservative: 0
Best Local Similarity: 98.31% Mismatches: 1
Query Match: 14,66% Indels: 0
DB: 9 Gaps: 0

US-10-077-111-13 (1-384) x US-09-864-761-27687 (1-180)

QY 227 ArgArgThnGluHsiGlnLeuYsgInphThGluAspTpsErgLguIuValIser 246
Db 178 AGCGGACAGACATCAGCTGAGACATTTACCGAAGATTGTCAGAGGAGATGTCTCA 119
QY 247 ThTTTLeuCYaAlaGlnApleuLyGAspLeuValGlyIleheLySmetAsnAnlle 266
Db 118 ACATGCTTTGTCACAAAGATTTAAAGACTTGTGTAATTTTCAAGATGAATTAACATT 59
QY 267 AAsGGLyEgLuLeuLeuAsnLeuThrLyGluSerLeuAlaAspApleuLySile 285
Db 58 GATGGAAAGAACTGTTGAACTTTACAAAGAAAGCTGGCTGATGATTTGAAATTT 2

RESULT 2
```

```

US-09-864-761-30542/c
; Sequence 30542, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 30542
; LENGTH: 132
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC009307.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.76
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.73
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.78
; OTHER INFORMATION: EXPRESSED IN HBLA, SIGNAL = 0.89
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.75
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.68
; OTHER INFORMATION: EST HUMAN HIT: BE675766.1, EVALUATE 2.00e-68
; OTHER INFORMATION: NT HIT: X71133.1, EVALUATE 1.20e+00
; OTHER INFORMATION: SWISSPROT HIT: P54860, EVALUATE 1.00e+00
US-09-864-761-30542

Alignment Scores:
Pred. No.: 7,04e-16 Length: 132
Score: 210.00 Matches: 43
```

Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 10.26%
DB: 9
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-10-077-111-13 (1-384) x US-09-864-761-30542 (1-132)

QY 290 LeuArgSerIysValLeuArgIysIleGluGluLeuArgThrIysValIysSerLeuSer 309
DB 130 CTGGGTGTAAGCGTGAAGGAAATGGAAGCTCAGACCAAGCTTAATCCCTTCT 71

QY 310 SerGlyIleProArgGluPheIleCysProIleThrArgGluLeuMetIysAspProVal 329
DB 70 TCAGGAATTCCTGATGATTTATATGTCACATACTAGAACTTAAGAAAGATCCGCTC 11

QY 330 IleAlaSer 332
DB 10 ATCGCATCA 2

RESULT 3

US-10-106-698-3381
; Sequence 3381, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: P4005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 3381
; LENGTH: 92
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45)..(45)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (78)..(78)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (86)..(86)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (89)..(89)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (91)..(91)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-3381

Alignment Scores:

Pred. No.: 5.64e-09 Length: 92
Score: 150.00 Matches: 26
Percent Similarity: 92.86% Conservative: 0
Best Local Similarity: 92.86% Mismatches: 2
Query Match: 7.33% Indels: 0
DB: 15 Gaps: 0

US-10-077-111-13 (1-384) x US-10-106-698-3381 (1-92)

QY 161 GlyAspLeuThrValITTPAspAspIysMetArgCysIleuHisSerGluIysAlaHisAsp 180
DB 1 GGTGATTTAAACAGTGTGGATGATTAATAATGAGTGTCTGTCATAGGAAAAAGACATGAT 60

QY 181 LeuGlyIleThrCysCysAspPhe 188
DB 61 CTGGAAATTAACCTGCTGATGATTT 84

RESULT 4

US-10-029-386-26571
; Sequence 26571, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 26571
; LENGTH: 174
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR3.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
; OTHER INFORMATION: SWISSPROT HIT: O60907, EVALUATE 3.00e-29
; OTHER INFORMATION: NT HIT: g14764641, EVALUATE 7.00e-90
; OTHER INFORMATION: EST_HUMAN HIT: BE465313.1, EVALUATE 1.00e-89
US-10-029-386-26571

Alignment Scores:

Pred. No.: 0.0841 Length: 174
Score: 93.50 Matches: 23
Percent Similarity: 49.12% Conservative: 5
Best Local Similarity: 40.35% Mismatches: 28
Query Match: 4.57% Indels: 1
DB: 16 Gaps: 1

US-10-077-111-13 (1-384) x US-10-029-386-26571 (1-174)

QY 71 SerCysSerThrAspGlyThrThrValIleuThrAsnThrGluAsnGlyGlnMetLeuAla 90

DB 3 AGTCTCTGTTGATTGATTCTACGTCGACCTGTGGACATTAAGACAGCGCTCTGCACCCAC 62

QY 91 ValMetGluGlnProSerGlySerProValArgValCysGlnPheSerProAspSerThr 110

DB 63 ACGCTCAGCAAGAGCATGAG---GAGCCTGTCTATAGCGTAGCTTTCAGCCCTGATGGAG 119

QY 111 CysLeuAlaSerGlyAlaAlaAspGlyThrValValIleuThrAsnAlaGln 127

DB 120 TACTTGGCCAGTGGATCTTCGACAGTGGCTCATCTTCGAACTACTCAG 170

RESULT 5

US-10-029-386-18677
; Sequence 18677, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 18677
; LENGTH: 150
; TYPE: DNA

```

ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC022073.13
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.5
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.8
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.5
OTHER INFORMATION: SWISSPROT HIT: Q92122, EVALUE 7.00e-26
OTHER INFORMATION: NT HIT: g45005931, EVALUE 3.00e-79
OTHER INFORMATION: EST_HUMAN HIT: BE892183.1, EVALUE 5.00e-79
US-10-029-386-18677

Alignment Scores:
Pred. No.:      0.396          length:      150
Score:         87.00          Matches:     18
Percent Similarity: 56.76%       Conservative: 3
Best Local Similarity: 48.65%    Mismatches: 16
Query Match:   4.25%           Indels:    0
DB:            16             Gaps:     0

US-10-077-111-13 (1-384) x US-10-029-386-18677 (1-150)
Qy      51  LyvphenHsthrTyxAlavAlHisCyvScySpheSerProsergiYhiIsileLeula 70
        |||  |||  ::|||::|||  |||  |||  |||  |||  |||  |||  |||
Db      9  AAGGACACTTTGGTCCTATTCACCTGTGAGACTTTAAGTCTGATGAGAAGACTTATGCC 68

Qy      71  SerCysSerThrAspGlyThrhValLeuTrpAsnThrGlubngIylgIn 87
        |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      69  AGTGCTTCAGAAAGATGGAACATTGAGACTTGCACCAACTGTGTAGGAAAA 119

RESULT 6
US-09-864-761-29419
? Sequence 29419, Application US/09864761
? Patent No. US20020048763A1
? GENERAL INFORMATION:
? APPLICANT: Penn. Sharron G.
? APPLICANT: Rank, David R.
? APPLICANT: Hanzel, David K.
? APPLICANT: Chen, Wensheng
? TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
? FILE REFERENCE: Aecmca-X-1
? CURRENT FILING DATE: 2001-05-23
? PRIOR APPLICATION NUMBER: US 60/180,312
? PRIOR FILING DATE: 2000-02-04
? PRIOR APPLICATION NUMBER: US 60/207,456
? PRIOR FILING DATE: 2000-05-26
? PRIOR APPLICATION NUMBER: US 09/632,366
? PRIOR FILING DATE: 2000-08-03
? PRIOR APPLICATION NUMBER: GB 24263.6
? PRIOR FILING DATE: 2000-10-04
? PRIOR APPLICATION NUMBER: US 60/236,359
? PRIOR FILING DATE: 2000-09-27
? PRIOR APPLICATION NUMBER: PCT/US01/00666
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00667
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00664
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00669
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00665
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00668
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00663
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00662
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00661
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00670
```

```

PREDICTION OF SIGNALING PATHWAY GENE EXPRESSION
US-10-077-111-13 (1-384) x US-09-864-761-29419 (1-173)

Alignment Scores:
Pred. No.:      1.96      Length:      173
Score:          82.00     Matches:      18
Percent Similarity: 55.56% Conservative: 12
Best Local Similarity: 33.33% Mismatches:    22
Query Match:    4.01% Indels:        2
DB:             9 Gaps:         1

US-10-077-111-13 (1-384) x US-09-864-761-29419 (1-173)

Qy 36 LeuTyrSerLeuArgAspRheThrGluLeuProHisSerProLeuYsrPheHisThrTyx 55
Db 7 CTTTGGAATATGGATAAATACACCATGATG-----CGGAATAAGAGAAGCATCACCAT 60
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 56 AlavaHtScyArCywCysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAsp 75
Db 61 GAATGGTAGCTACTGTACTCTTTCTCTGATGGAGCAATTGAGCTACTGCATCTTATGAT 120
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 76 GlyThrValLeuItpAenThrgIuaSngIylMeLcu 89
Db 121 ACTGAGATATATATCTGGATCCACATTAATGAGACATTTCTG 162
   ::|||::| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 7
US-10-305-720-417
; Sequence 417, Application US/10305720
; Publication No. US20040010136A1
GENERAL INFORMATION:
APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.
TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expression
FILE REFERENCE: PA-0002-1 CON
CURRENT APPLICATION NUMBER: US/10/305,720
PRIOR FILING DATE: 2002-11-26
PRIOR APPLICATION NUMBER: 09/016,434
NUMBER OF SEQ ID NOS: 1490
SOFTWARE: PERL Program
SEQ ID NO 417
LENGTH: 169
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No. US20040010136A1 2137838
US-10-305-720-417

Alignment Scores:
Pred. No.:      3.75      Length:      169

```


Score: 79.50 Matches: 19
Percent Similarity: 53.57% Conservative: 11
Best Local Similarity: 33.93% Mismatches: 25
Query Match: 3.88% Indels: 1
DB: 17 Gaps: 1

US-10-077-111-13 (1-384) x US-10-305-720-417 (1-169)

QY 69 Leu1a1aSerCysSerThrAspGlyThrValLeuTrpAsnThrGluAsnGlyGluMet 88
DB 3 CTCATCAGAGCTTCTTCAGATGATGATCCTTAAGATCTTGACCTTAAAGAAAGGCTC 62
QY 89 Leu1a1aValMetGluGlnProSerGlySerProValArgValCysGlnPheSerProAsp 108
DB 63 ATCATATCACTTCAAGACATACGGGA---CCTGCTTACTGTTTCAAAAAGT 119
QY 109 SerThrCysLeu1a1aSerGlyAla1a1aSerGlyThrValValLeuTrp 124
DB 120 GGAGAGCTATTGTCATCAGAGGTGCAGACACACAGAGCTTATTATG 167

RESULT 8

US-10-029-386-16111/C
Sequence 16111, Application US/10029386
Publication No. US20030194704A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David K.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
FILE REFERENCE: AROMICA-X-2
CURRENT APPLICATION NUMBER: US/10/029,386
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 16111
LENGTH: 155
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO CHR14.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.1
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.9
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.2
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.7
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.5
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.4
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.8
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.2
OTHER INFORMATION: SWISSPROT HIT: P38262, EVALUE 5.90e-02
OTHER INFORMATION: NT HIT: gi13649630, EVALUE 8.00e-74
OTHER INFORMATION: EST_HUMAN HIT: BE739840.1, EVALUE 1.00e-73
US-10-029-386-16111

Alignment Scores:

Pred. No.: 4.3 Length: 155
Score: 78.50 Matches: 19
Percent Similarity: 40.58% Conservative: 9
Best Local Similarity: 27.54% Mismatches: 18
Query Match: 3.83% Indels: 24
DB: 16 Gaps: 1

US-10-077-111-13 (1-384) x US-10-029-386-16111 (1-155)

QY 107 ProAspSerThrCysLeu1a1aSerGlyAla1a1aSerGlyThrValValLeuTrpAsnAla 126
DB 144 CCAGATGACATGATATACCAAGAGGTCTGACCCGCTGCTGCGCTTG----- 91
QY 127 GlnSerThrCysLeu1a1aSerGlyAla1a1aSerGlyThrValValLeuTrpAsnAla 146
DB 91 ----- 91
QY 147 PheSerProAsnGlySerPhePheValThrGlySerSerCysGlyAspLeuThrValTrp 166

DB 90 ---GCACAGATGTTCCATGACAGATATCTGGAATCAAGCTGGGAACTATCTTGTGG 34
QY 167 AspaAspSerMetArgCysLeuHisSer 175
DB 33 CAGGAAGCTAAGGC-TGTGGCCACAGC 8

RESULT 9

US-09-998-598-205/C
Sequence 205, Application US/09998598
Patent No. US20020150922A1
GENERAL INFORMATION:
APPLICANT: Stolk, John A.
APPLICANT: Xu, Jiangchun
APPLICANT: Chenault, Ruth A.
APPLICANT: Meagher, Madelein Joy
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.561
CURRENT APPLICATION NUMBER: US/09/998,598
CURRENT FILING DATE: 2001-11-16
NUMBER OF SEQ ID NOS: 2606
SOFTWARE: Corixa Invention Disclosure Database
SEQ ID NO 205
LENGTH: 197
TYPE: DNA
ORGANISM: Homo sapiens
US-09-998-598-205

Alignment Scores:

Pred. No.: 32.6 Length: 197
Score: 72.50 Matches: 21
Percent Similarity: 44.62% Conservative: 8
Best Local Similarity: 32.31% Mismatches: 17
Query Match: 3.54% Indels: 19
DB: 9 Gaps: 3

US-10-077-111-13 (1-384) x US-09-998-598-205 (1-197)

QY 31 AspLeuThr11a1aGlu-----TyrSerLeuArgAspPhe 42
DB 194 GATTAACCATCAAGATATGGAATACCTGGGTGTGCAAAATGACCTGTCAGAT--- 138
QY 43 ThrGluLeuProHisSerProLeuLysPheHisThrTyrAlaValHisCysCysPhe 62
DB 137 -----GAGAGCCACTCAGAGTGGGTCTTGTGTCGCTTC 102
QY 63 SerPro-----SerGlyHis1a1aSerCysSerThrAspGlyThrValLeu 80
DB 101 TCGCCCAACAGCAGCAACCTATATCATGCTCTCTGCTGGCGAGCAAGCTGCAAGGTA 42
QY 81 TrpAsnThrGluAsn 85
DB 41 TGAACCTGGCTAAC 27

RESULT 10

US-09-998-598-2327
Sequence 2327, Application US/09998598
Patent No. US20020150922A1
GENERAL INFORMATION:
APPLICANT: Stolk, John A.
APPLICANT: Xu, Jiangchun
APPLICANT: Chenault, Ruth A.
APPLICANT: Meagher, Madelein Joy
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.561
CURRENT APPLICATION NUMBER: US/09/998,598
CURRENT FILING DATE: 2001-11-16
NUMBER OF SEQ ID NOS: 2606
SOFTWARE: Corixa Invention Disclosure Database
SEQ ID NO 2327
LENGTH: 197

```
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-998-598-2327

Alignment Scores:
Pred. No.: 32.6      Length: 197
Score: 72.50      Matches: 21
Percent Similarity: 44.62%      Conservative: 8
Best Local Similarity: 32.31%      Mismatches: 17
Query Match: 3.54%      Indels: 19
DB: 9      Gaps: 3

US-10-077-111-13 (1-384) x US-09-998-598-2327 (1-197)

QY 31 AppLyseThrIleArgLeu-----TyrSerLeuArgAspPhe 42
   |||||
Db 4 GATTAACCATCAAGCATGGAATACCTGGGTGTGCAAAATCACTGCGAGAT--- 60
QY 43 ThrGluLeuProHisSerProLeuLysPheHisThrTyraIaValHisCysCysPhe 62
   |||||
Db 61 -----GAGAGCCACTCAGAGTGGGTGTCTGTGTCGCCCTTC 96
QY 63 SerPro-----SerGlyHisIleLeuAlaSerCysSerThrIleArgIleThrValLeu 80
   |||||
Db 97 TCGCCCAACAGCAGCAACCTATCATCTGCTCTGTGGCTGGACAGACTGTCAAGTA 156
QY 81 TrpAsnThrGluAsn 85
   |||||
Db 157 TGGAACTGGCTAAC 171

RESULT 11
US-10-425-115-24663
/ Sequence 24663, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovacic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53222)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ NUMBER OF SEQ ID NOS: 2003-04-28
/ SEQ ID NO 24663
/ LENGTH: 187
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MRT4577_1224C.1
US-10-425-115-24663

Alignment Scores:
Pred. No.: 34.5      Length: 187
Score: 72.00      Matches: 14
Percent Similarity: 52.94%      Conservative: 4
Best Local Similarity: 41.18%      Mismatches: 16
Query Match: 3.52%      Indels: 0
DB: 18      Gaps: 0

US-10-077-111-13 (1-384) x US-10-425-115-24663 (1-187)

QY 345 TrpIleSerIysIysArgThrSerProMetThrAsnLeuValLeuProSerAlaVal 364
   |||||
Db 2 TGGTTCGGCCCGCCGCAAGTGTCTGCCCAAGCCCGCAGGAGCTGACTAATCTGGAG 61
QY 365 LeuThrProAsnArgThrLeuLysMetAlaIleAsnArgTrp 378
   |||||
Db 62 CTGTGTGCCCAAGAGCGCTCAAGAACTGATCTCGCGGTGG 103

RESULT 12
US-10-424-599-124934
```

```
/ Sequence 124934, Application US/10424599
/ Publication No. US20040031072A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovacic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53223)B
/ CURRENT APPLICATION NUMBER: US/10/424,599
/ NUMBER OF SEQ ID NOS: 285684
/ SEQ ID NO 124934
/ LENGTH: 169
/ TYPE: DNA
/ ORGANISM: Glycine max
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT3847_83824C.1
US-10-424-599-124934

Alignment Scores:
Pred. No.: 33.7      Length: 169
Score: 71.50      Matches: 15
Percent Similarity: 50.00%      Conservative: 9
Best Local Similarity: 31.25%      Mismatches: 17
Query Match: 3.49%      Indels: 7
DB: 17      Gaps: 1

US-10-077-111-13 (1-384) x US-10-424-599-124934 (1-169)

QY 125 AsnAlaGlnSerTyrlsLeuTyraArgCysGlySerValIysAspGlySerLeuAlaIa 144
   |||||
Db 23 AATGCAATGCCAACAAATCTTACGAGGTATCCCAACACCTCTGATGCAATATCTAGT 82
QY 145 CysAlaPheSerProAsnGlySerPheIleValThrGlySerCysGlyAsnLeuThr 164
   |||||
Db 83 ATTAGGTCAGTCCCAAGCAACATGTCTTCGTCACCT-----ACT 121

QY 165 ValTrpAspAspLysMetArgCys 172
   |||||
Db 122 TCATGGGACACCAAGTTGGGTGT 145

RESULT 13
US-10-029-386-25501/c
/ Sequence 25501, Application US/10029386
/ Publication No. US20030194704A1
/ GENERAL INFORMATION:
/ APPLICANT: Penn, Sharon G.
/ APPLICANT: Hanzel, David R.
/ APPLICANT: Hanzel, David K.
/ TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
/ FILE REFERENCE: AEOMICA-X-2
/ CURRENT APPLICATION NUMBER: US/10/029,386
/ NUMBER OF SEQ ID NOS: 2001-12-20
/ SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
/ SEQ ID NO 25501
/ LENGTH: 131
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO CHR19.3
/ OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.2
/ OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.6
/ OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
/ OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8
/ OTHER INFORMATION: SWISSPROT HIT: P90587, EVALUOE 4.00e-04
/ OTHER INFORMATION: EST HUMAN HIT: B1560259.1, EVALUOE 5.00e-23
/ OTHER INFORMATION: NT HIT: AF072879.1, EVALUOE 4.40e-01
US-10-029-386-25501
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: February 5, 2005, 11:55:52 ; Search time 694 Seconds
(without alignments)
3184.733 Million cell updates/sec

Title: US-10-077-111-13
Perfect score: 2047
Sequence: 1 MVXLIHTLADHGDDVNCACF.....LTPNRLTKMAINWLETHQK 384

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapext 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4313806 segs, 2877871033 residues
Total number of hits satisfying chosen parameters: 8627612

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=x1h
-O=/cg2_1/USPTO.spool/US10077111/runat_04022005_121054_6859/app.query.fasta.1.583
-DB=Published Applications NA -QFWT=fastap -SUFFIX=p2n.rmpb -MINMATCH=0.1
-LOOPCU=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=-1 -MATRIX=bl6sum62
-TRANS=human40.cdt -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MALEN=2000000000 -USER=US10077111 @CGN_1_1_628@runat_04022005_121054_6859
-NCPU=6 -ICPU=3 -NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:*

1: /cg2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cg2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cg2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cg2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cg2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cg2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cg2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cg2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cg2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
10: /cg2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cg2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cg2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cg2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cg2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
15: /cg2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cg2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
17: /cg2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
18: /cg2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
19: /cg2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
20: /cg2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
21: /cg2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
22: /cg2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2047	100.0	1272	13	US-10-077-111-12
2	1984	96.9	1553	13	US-10-077-111-3
3	1984	96.9	1818	13	US-10-077-111-1
4	1984	96.9	1996	18	US-10-357-930-23179
5	1984	96.9	1996	18	US-10-357-930-29046
6	1978	96.6	1773	16	US-10-287-218-40
7	1978	96.6	1773	18	US-10-474-291-40
8	1978	96.6	1811	17	US-10-104-047-1689
9	1978	96.6	1817	10	US-09-971-392-207
10	1978	96.6	1844	15	US-10-037-720-768
11	1978	96.6	1844	17	US-10-117-722-768
12	1958	95.7	1908	13	US-10-077-111-14
13	1616.5	79.0	1901	13	US-10-077-111-6
14	709	34.6	630	13	US-10-077-111-5
15	627.5	30.7	446	18	US-10-357-930-5029
16	626.5	30.6	409	18	US-10-357-930-14198
17	621.5	30.4	441	18	US-10-357-930-35319
18	621.5	30.4	441	18	US-10-357-930-44154
19	597	29.2	366	9	US-09-864-761-10657
20	532	26.0	297	9	US-09-864-761-27306
21	352	17.2	520	13	US-10-077-111-8
22	308.5	15.1	466	9	US-09-864-761-11052
23	300	14.7	180	9	US-09-864-761-27687
24	264.5	12.9	837	17	US-10-369-493-43585
25	264	12.9	3015	17	US-10-369-493-43529
26	263.5	12.9	1593	17	US-10-369-493-36600
27	259.5	12.7	3025	9	US-09-764-853-29
28	259	12.7	3082	17	US-10-451-207-18
29	256.5	12.5	3166	17	US-10-451-207-13
30	251	12.3	1137	17	US-10-369-493-31840
31	250	12.2	1741	17	US-10-369-493-42580
32	246.5	12.0	1830	17	US-10-369-493-43872
33	245	12.0	840	17	US-10-369-493-43586
34	244.5	11.9	3420	17	US-10-369-493-42740
35	242.5	11.8	1965	17	US-10-369-493-42753
36	242.5	11.8	3408	17	US-10-369-493-42733
37	236.5	11.6	3358	17	US-10-369-493-42532
38	232.5	11.4	3465	17	US-10-369-493-42536
39	231.5	11.3	1818	17	US-10-369-493-42594
40	230	11.2	852	17	US-10-369-493-42751
41	230	11.2	2768	17	US-10-424-599-115833
42	229.5	11.2	852	17	US-10-369-493-42703
43	229	11.2	480	9	US-09-864-761-13978
44	229	11.2	1434	17	US-10-369-493-42592
45	227	11.1	7042	9	US-09-876-667-1

ALIGNMENTS

RESULT 1
US-10-077-111-12
: Sequence 12, Application US/10077111
: Publication No. US20020187492A1
: GENERAL INFORMATION:
: APPLICANT: Toddertud, C. Gordon
: APPLICANT: Flinger, Joshua N.
: APPLICANT: Rillema, Jijia
: TITLE OF INVENTION: TBA
: FILE REFERENCE: 3053-411AUS2
: CURRENT APPLICATION NUMBER: US/10/077, 111
: PRIOR FILING DATE: 2002-02-15
: PRIOR APPLICATION NUMBER: 60/294,181
: PRIOR FILING DATE: 2001-05-29
: PRIOR APPLICATION NUMBER: 60/269,366
: NUMBER OF SEQ ID NOS: 25
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 12
: LENGTH: 1272

```
/ TYPE: DNA
/ ORGANISM: HUMAN
/ FEATURE:
/ OTHER INFORMATION: RET 16.2 splice variant
US-10-077-111-12

Alignment Scores:
Pred. No.: 2,74e-233 Length: 1272
Score: 2047.00 Matches: 384
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-077-111-13 (1-384) x US-10-077-111-12 (1-1272)

QY 1 MetValLyLeuLlLhIstHrLeuAlaAphHlSGlYAspAspValaenCYsCyAlaPhe 20
DB 111 ATGGTAAACTGATTACACATTAGCTGATCATGGTACGATGCAAGTGTGCTGCTTC 170
QY 21 SerPheSerLeuLeuAlaThrCYsSerLeuAspLyThrIleArgLeuTYrSerLeuArg 40
DB 171 TCTTTTCCCTCTTGCTACTTGTCTCTTGACAAACATTCGCTGTACTGTTACGT 230
QY 41 AspPheThrGluleuProHlSerProLeuLyPheHlThrTyAlaValHlScyCyS 60
DB 231 GACTTACTGAACTGCGCACATCTCCATTGAACTTTCATCTCATGCTGTCCACTGCTG 290
QY 61 CyAPheSerProSerGlyHlIleLeuAlaSerCYsSerThrAspGlyThrThrValLeu 80
DB 291 TGTTCCTCCCTTCAGACATATTGGCATCGTGTTCACAAATGGTACACTGTCTCTA 350
QY 81 TrpAsnThrGlulaenGlyInMeCleuAlaValMeGlunInProSerGlySerProVal 100
DB 351 TGGAAATACGAAATAGACAGATGCTGCGCATGATGAAACGCTAGTGCACCTCTGTC 410
QY 101 ArgValCYsGlnPheSerProAspSerThrCYsLeuAlaSerGlyAlaIleAspGlyThr 120
DB 411 AGGGTTGGCCAGTTTCCCGACACTCCAGCTGTTGGCATCAGGGGCGAGCTGATGAACT 470
QY 121 ValValLeuTrpAsnAlaGlnSerTYrLyLeuTYrArgCYsGlySerValIlyAspGly 140
DB 471 GTGGTTTGGGATGCACAGTATCAATTAATTAAGATGTGTGTAAAGATGCG 510
QY 141 SerLeuAlaAlaCYsAlaPheSerProAsnGlySerPhePheValThrGlySerSerCYs 160
DB 531 TCTTGGCGGCAATGCTATTTCTCTTAATGAAAGCTCTTGTCTACTGCTCTCATGT 590
QY 161 GlyAspLeuThrValITrpAspAspLyMeLArgCYsLeuHlSerGlyValaAlaHlAsp 180
DB 591 GGTGATTTAAACAGTGTGGGATGATAAATAGAGGTGTCTGCATAGTAAAGAACATGAT 650
QY 181 LeuGlyIleThrCYsCyAspPheSerSerGlnProValSerAspGlyGluGlnGlyLeu 200
DB 651 CTTCGAATTAAGTGTGCGCATTTTCTTCACACAGTTCATGATGAGAACAGGCTTT 710
QY 201 GlnPhePheArgLeuAlaSerCYsGlyGlnAspCYsGlnValIlySerPleValSer 220
DB 711 CAGTTTTCGACTGCGCATCATGTGTCAGGATTCGCAAGTTCGAAATTTGATGTTTCT 770
QY 221 PheThrHlIleLeuAlaArgArgThrGlnHlGlnLeuLySglnPheThrGluAspTrp 240
DB 771 TTATACCATATCTTACCAAGCGCACAGAACATCAGTGAAGCAATTTACCGAAGATTGG 830
QY 241 SerGluGluValIySerThrTrpLeuCYsAlaGlnAspLeuLyAspLeuValIyIle 260
DB 831 TCAGAGAGAGTCTGTCAACATGCTTTGTGCAACAGATTAAAGATCTTGTGTGGATT 890
QY 261 PheLyMeLAsnAsnIleAspGlyLySglnLeuLeuAsnLeuThrLySglnSerLeuAla 280
DB 891 TTCGAATGAATTAACATTGATGAGAAAGAACTGTTGAATTTTCAAAAGAAAGTCTGGCT 950
QY 281 AspAspLeuLyIleGlnSerLeuGlyLeuArgSerLyValLeuArgLyIleGlnGlu 300
```

```
DB 951 GATGATTTGAAATGGAATCTCTAGACATCGCTAGTAAAGTCTGAGGAAATTTGAAGAG 1010
QY 301 LeuArgThrLyLeuValySerLeuSerSerGlyIleProAspGluPheIleCYsProIle 320
DB 1011 CTCAGACCAAGGTTTAAATCCCTTTCTTCAGGAATCTCTGATGAAATTTATATGTCCAAAT 1070
QY 321 ThrArgGluMeLysAspProValIleAlaSerAspGlyTYrSerTYrGluLySgln 340
DB 1071 ACTAGAGAACTTATGAAAGATCCGGTCATCGATCGATGATGGCTATTATCATGAAAGAA 1130
QY 341 AlaMeGlunAsnTrpIleSerLySglnValArgThrSerProMetThrAsnLeuValLeu 360
DB 1131 GCATGGAATAATGGATTCAGCAAAAGAAACGATCAAGTCCCATGCAATCTGTGTTCTT 1190
QY 361 ProSerAlaValLeuThrProAsnArgThrLeuLyMeLAlaIleAsnArgTrpLeuGlu 380
DB 1191 CTTTCAGCGGTACTTACCAAAATAGACTCTGAAATGGCCATTCATATGATGGCTGGAG 1250
QY 381 ThrHlGlnLyS 384
DB 1251 ACACACCAAAAG 1262

RESULT 2
US-10-077-111-3
/ Sequence 3, Application US/10077111
/ Publication No. US20020187492A1
/ GENERAL INFORMATION:
/ APPLICANT: Toddertud, C. Gordon
/ APPLICANT: Ringer, Joshua N.
/ APPLICANT: Rillema, J11
/ TITLE OF INVENTION: TBA
/ FILE REFERENCE: 3053-4114US2
/ CURRENT APPLICATION NUMBER: US/10/077,111
/ PRIOR FILING DATE: 2002-02-15
/ PRIOR FILING DATE: 2001-05-29
/ PRIOR APPLICATION NUMBER: 60/294,181
/ PRIOR FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 25
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 3
/ LENGTH: 1553
/ TYPE: DNA
/ ORGANISM: HUMAN
US-10-077-111-3

Alignment Scores:
Pred. No.: 1,22e-225 Length: 1553
Score: 1984.00 Matches: 383
Percent Similarity: 80.46% Conservative: 0
Best Local Similarity: 96.46% Mismatches: 1
Query Match: 96.92% Indels: 92
DB: 13 Gaps: 1

US-10-077-111-13 (1-384) x US-10-077-111-3 (1-1553)

QY 1 MetValLyLeuLlLhIstHrLeuAlaAphHlSGlYAspAspValaenCYsCyAlaPhe 20
DB 111 ATGGTAAACTGATTACACATTAGCTGATCATGGTACGATGCAAGTGTGCTGCTTC 170
QY 21 SerPheSerLeuLeuAlaThrCYsSerLeuAspLyThrIleArgLeuTYrSerLeuArg 40
DB 171 TCTTTTCCCTCTTGCTACTTGTCTCTTGACAAACATTCGCTGTACTGTTACGT 230
QY 41 AspPheThrGluleuProHlSerProLeuLyPheHlThrTyAlaValHlScyCyS 60
DB 231 GACTTACTGAACTGCGCACATCTCCATTGAACTTTCATATCTCATGCTGTCCACTGCTG 290
QY 61 CyAPheSerProSerGlyHlIleLeuAlaSerCYsSerThrAspGlyThrThrValLeu 80
DB 291 TGTTCCTCCCTTCAGACATATTGGCATCGTGTTCACAAAGATGATGCAAGTGTCTCTA 350
```

QY 81 TTPASnThrgluansnglyglmetleuAlaValmetgluInProserglySerProval 100
DB 351 TGGAAATACGAAAAATGACAGATGCTGGCAGTGTGAGAACAGCTTACTGACAGCCCTGTG 410
QY 101 ArgValCysglInheserProaspsSerThrCysleuAlaSerGlyValaaIaaspGlyThr 120
DB 411 AGGGTTGGCCAGTTTCCCAAGACTCCAGCTGTTGGCATCAGGGGAGCGTGAAGACT 470
QY 121 ValValleuThrpsanaIaInserTyrLysleuTyrArgCysglSerValIysaspGly 140
DB 471 GTGGTTTGTGGAATGACAGTCAATCAAAATTAATATGATGTGTAGTGTAAAGATGCG 530
QY 141 SerleuAlaAlaCysAlaPheSerProasnglySerPhePheValThrGlySerSerCys 160
DB 531 TCCTGGCGGCATGTGATTTCTCCAAATGGAAGCTTCTTGTGCACTGCTCCTCATGT 590
QY 161 GlyAspLeuThrValTTPAspAspLysMetArgCysleuHisSerGlyLysAlaHisAsp 180
DB 591 GGTGATTTAACAGTGTGGATGATMAAATAGGTGTGATGTGAAAAAGCACATGAT 650
QY 181 LeuGlyIleThrCysAspPheSerSerGlnProValSerAspGlyGluGlnGlyLeu 200
DB 651 CTGGAAATTAAGCTGCTGCGATTTTCTTCAAGCCAGTTTCTGATGAGAACAGATCTT 710
QY 201 GlnPhePheArgleuAlaSerCysGlyGlnAspCysGlnValIysIleTyrPileValSer 220
DB 711 CAGTTTCTTGCATGGCATCATGTGTGACAGATGGCAAGTCAAAATTTGGATTGTTCT 770
QY 221 PheThrHisIleLeu----- 225
DB 771 TTTACCCATATCTTAGGTTTGAATTAATAATAAGTACACTGAGTGGGCACTGTGCT 830
QY 225 ----- 225
DB 831 CCTGTTCTGCGCTTGTCTTTTCCCGTGAAGGCGACATGCTAGTCTCAGGGTCAATGGAT 890
QY 225 ----- 225
DB 891 AAGTGTGTCAATGATATGATACTAATACTGAGAAATATACTTCACACATGACTCAGCAC 950
QY 225 ----- 225
DB 951 ACCAGTATGTCAAACTTGTGCTTTTGCACCTAATACCTTTTACTTGTCTACTGATTCA 1010
QY 226 -----AlaArgArg 228
DB 1011 ATGGAACAAAACAGTAAACATCTGGCAATTTGACCTGGAAACACTTTGCCAAGCAAGCGC 1070
QY 229 ThrGluHisglInleuLysglInPheThrGluAspTyrSerGlyGluValIleSerThrTyr 248
DB 1071 ACAGAACATCAGCTGAAGCAATTTACCGAAGATTTGGTCAGAGAGATGTCTCAACATGG 1130
QY 249 LeuCysAlaGlnAspLeuLysAspLeuValGlyIlePheLysMetAsnAsnIleAspGly 268
DB 1131 CTGTGTCACAAAGATTTAAAGATCTGTGTATTTTCAAGATGAATTAACATTGATGGA 1190
QY 269 LysGluLeuLeuAsnLeuThrLysGluSerleuAlaAspAspLeuLysIleGluSerLeu 288
DB 1191 AAAGAACTGTGTGAATCTTCAAAAAGAAAGTCTGGCTGATATTTGAAAATTTGAATCTCTA 1250
QY 289 GlyLeuArgSerLysValleuArgLysIleGluGluLeuArgThrLysValIysSerLeu 308
DB 1251 GGAAGTGTGTAAGTGAAGCTGAGGAAATTTGAAGAGCTCAGACCAAGGTTAAATCCCTT 1310
QY 309 SerSerGlyIleProAspGluPheIleCysProIleThrArgGluLeuMetLysAspPro 328
DB 1311 TCTTCAGGAATTCCTGATGATTAATATATGTCATAATACAGAGAACTTATGAAGAATCCG 1370
QY 329 ValIleleIaSerAspGlyTyrSerTyrGluLysGluAlaMetGluAsnTyrPileSerLys 348
DB 1371 GTATTCGCATCAAGTGGCTATTCATATGAAAAGAACCAATGGAATAATTTGATCAGCAA 1430
QY 349 LysLysArgThrSerProMetThrAsnLeuValleuProSerAlaValleuThrProAsn 368

DB 1431 AAGAAACGTACAGATCCCATGACAAATCTTGTCTTCCACGGGTACTTACACCAAT 1490
QY 369 ArgThrLeuLysMetAlaIleAsnArgTyrPheGluThrHisGlnLys 384
DB 1491 AGGACTCTGAAAATGGCCATCAATATAGATGGCTGGAGACACACCAANG 1538
RESULT 3
US-10-077-111-1
; Sequence 1, Application US/10077111
; Publication No. US20020187492A1
; GENERAL INFORMATION:
; APPLICANT: Toddertud, C. Gordon
; APPLICANT: Finger, Joshua N.
; APPLICANT: Rillema, Jill
; TITLE OR INVENTION: TBA
; FILE REFERENCE: 3053-4114US2
; CURRENT APPLICATION NUMBER: US/10/077,111
; CURRENT FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 60/294,181
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: 60/269,366
; PRIOR FILING DATE: 2001-02-16
; SOFTWARE: PatentIn Ver. 2.1
; NUMBER OF SEQ ID NOS: 25
; SEQ ID NO 1
; LENGTH: 1818
; TYPE: DNA
; ORGANISM: HUMAN
US-10-077-111-1
Alignment Scores:
Pred. No.: 1 566-225 Length: 1818
Score: 1984.00 Matches: 383
Percent Similarity: 80.46% Conservative: 0
Best Local Similarity: 80.46% Mismatches: 1
Query Match: 96.92% Indels: 92
DB: 13 Gaps: 1
US-10-077-111-13 (1-384) x US-10-077-111-1 (1-1818)
QY 1 MetValLysleuIleHisThrleuAlaAspHisGlyAspAspValAsnCysCysAlaPhe 20
DB 148 ATGGTGAACCTGATTCACACATTAAGTGAATCATGTGATGATGATGATGATGATGATGATG 207
QY 21 SerPheSerleuLeuAlaThrCysSerleuAspLysThrIleArgLeuTyrSerleuArg 40
DB 208 TCCCTTTCCTCTTGGCTACTGCTCTTGGACAAAACAATTCGCTGATCTGTTACGT 267
QY 41 AspPheThrGluLeuProHisSerProleuLysPheHisThrTyrAlaValHisCysCys 60
DB 268 GACTTACTGAACTGACCACTTCTCCATTTGAAGTTTCATACCTATCTGTCACCTGCTGC 327
QY 61 CysPheSerProSerGlyHisIleleuAlaSerCysSerThrAspGlyThrValIleu 80
DB 328 TGTTTCTCCCTTCAGACATATTTTGGCATCGTGTTCACAGATGCTACACTGTCTCA 387
QY 81 TTPASnThrgluansnglyglmetleuAlaValmetgluInProserglySerProval 100
DB 388 TGGAAATACGAAAATGACAGATGTGGCAGTATGAAACAGCTGTGCGACCCCTGTG 447
QY 101 ArgValCysglInheserProaspsSerThrCysleuAlaSerGlyValaaIaaspGlyThr 120
DB 448 AGGGTTGGCCAGTTTCCCAAGCTCCAGCTCAGTGTGTGACATCAGGGGAGCGTGAAGACT 507
QY 121 ValValleuThrpsanaIaInserTyrLysleuTyrArgCysglSerValIysaspGly 140
DB 508 GTGGTTTGTGGAATGACAGTCAATCAAAATTAATATGATGTGTAGTGTAAAGATGCG 567
QY 141 SerleuAlaAlaCysAlaPheSerProasnglySerPhePheValThrGlySerSerCys 160
DB 568 TCCTGGCGGCATGTGATTTCTCCAAATGGAAGCTTCTTGTGCACTGCTCCTCATGT 627

QY 161 G1yAaPLeuThrVal1rPaAaPlyMeTArCySeLeuH1SeSerG1uLySa1aH1aAaP 180
| | | | |
Db 628 GGTGATTTAAAGTGTGGGTGATGATTAAGTGTCTGTGATGTGATTAAGTGTGATGAT 687
| | | | |
QY 181 LeuG1y11eThrCySeCyAaPPhSeSerSerG1nProValSerAaPlyG1uG1nG1yLeu 200
| | | | |
Db 688 CTGGGAATTAACCTGCTGCGCATTTCTTCACAGCCAGTTCTGATGAGAAACAAGGCTT 747
| | | | |
QY 201 G1nPhaPheArGLeuH1aSeCySeG1yG1nAaPSeG1nVal1yS1eTTP11eVal1Ser 220
| | | | |
Db 748 CAGTTTTCGACGTGGCATATGTGTGACAGATTGCAACTCAAAATTGATTTGTTCTT 807
| | | | |
QY 221 PheThH1a1eLeu----- 225
| | | | |
Db 808 TTTAACCATATCTTAAGTTTGAATTAAATATAAGTACACTGAGTGGGCACTGTGCT 867
| | | | |
QY 225 ----- 225
| | | | |
Db 868 CCTGTTCTGGCTTGTGCTTTTCCATGATGGGCAAGTCTAGTCTCAGGGTCAGTGAT 927
| | | | |
QY 225 ----- 225
| | | | |
Db 928 AAGTGTCACTAGTATATGATTAATACTAGAGAAATTAATTAATTAATTAATTAATTA 987
| | | | |
QY 225 ----- 225
| | | | |
Db 988 ACCAGGATATGCACAACTTGCTGCTTTTGACACTAATACCTTTTACTGCTACTGTGCT 1047
| | | | |
QY 226 ----- 226
| | | | |
Db 1048 ATGACAAAACAGTGAACATCTGGCAATTGACCTGGAAACACTTGGCAAGAGGCGC 1107
| | | | |
QY 229 ThG1uH1eG1nLeuLySeG1nPhaThG1uAaPTrSeG1uG1uVal1y1SeThrTrP 248
| | | | |
Db 1108 ACGAACAATACGTGGAAGCAATTTACGAGATTTGGTCAAGAGATGTCTCAACTG 1167
| | | | |
QY 249 LeuCyAa1aG1nAaPLeuLySaPLeuVal1G1y11ePheLySeMeTArNaen11eAaPly 268
| | | | |
Db 1168 CTTTGTGACAAAGATTTAAAGATCTTGTGATTTTCAAGATGAATTAACATTTGATGA 1227
| | | | |
QY 269 LySeG1uLeuLySeThrLySeG1uSeSerLeuH1aAaPLeuLyS1eG1uSeSerLeu 288
| | | | |
Db 1228 AAGAACTGTGATCTTCAAAAAGAAAGTCTGCTGATGATTTGAAATTTGAATCTCTA 1287
| | | | |
QY 289 G1yLeuArSeSerLySeVal1euaRgLyS1eG1uLeuArThrLySVal1ySeSerLeu 308
| | | | |
Db 1288 GGACTGCGTAAAGTGTGAGAAATTTGAAGAGTCTGAGACCAAGGTTAAATCCCTT 1347
| | | | |
QY 309 SeSerG1y11ePProAaPlyuPhe1eCyPPro1eThrArG1uLeuWetLySaPPro 328
| | | | |
Db 1348 TCTTCAGGAATTCCTGATGAATTTATATGTCCAATACTAGAGACTTATGAAGAATCCG 1407
| | | | |
QY 329 Val11e1aSeArAaPlyTyTrSerTyRg1uLyG1uA1aMeG1uAaPTr11eSerLyS 348
| | | | |
Db 1408 GTCATGCAACAGATGCTATTCATATGAAAGCAATGAGCAATGAGATCAGAGAA 1467
| | | | |
QY 349 LySeLyArThrSePProMeTThrAaPLeuVal1euaProSeR1aVal1eThrProAa 368
| | | | |
Db 1468 AAGAAACGTACAGTCCCATGACCAATCTTCTTCTTCAGGGGACTTACCAAAAT 1527
| | | | |
QY 369 ArGThrLeuLySeMeTAr1eAaArGTrPLeuG1uThrH1eG1nLyS 384
| | | | |
Db 1528 AGGACTGTGAATAATGGCCATCAATAGATGCTGGAAGCACACCAAAAG 1575
| | | | |
RESULT 4
US-10-357-930-23179 ; Sequence 23179, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR

; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23179
; LENGTH: 1996
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1, 2, 3, 1987, 1988, 1989, 1990, 1991, 1992, 1993, 1994,
; LOCATION: 1995, 1996
; OTHER INFORMATION: n = A, T, C or G
US-10-357-930-23179
Alignment Scores:
Pred. No.: 1,81e-225 Length: 1996
Score: 1984.00 Matches: 383
Percent Similarity: 80.468 Conservative: 0
Best Local Similarity: 80.468 Mismatches: 1
Query Match: 96.924 Indels: 92
DB: Gaps: 1
US-10-077-111-13 (1-384) x US-10-357-930-23179 (1-1996)
QY 1 MetVal1yLeu11eH1eThrLeuH1aAaPPh1eG1yAaPArVal1aAaCySeVal1aPhe 20
| | | | |
Db 16 ATGTAAGAACTGATTCACACATTAAGTGTATCATGTGAGAGATGTCAACTGTGCTTC 75
| | | | |
QY 21 SePheSerLeuLeuH1aThrCySeSerLeuAaPlyThr11eArGLeuTyTrSerLeuAa 40
| | | | |
Db 76 TCTTTTCCCTCTTGCTGCTCTGCTCTTGGACAAACAAATTCGCTGTAATCTTACGT 135
| | | | |
QY 41 AaPPhaThG1uLeuPProH1aSePProLeuLySPha1eThrTyR1aVal1y1a1ySaCyS 60
| | | | |
Db 136 GACTTACTGAACTGCCACATTTCTCATTTGAAGTTTCATTAATCTTAATCTTCACTGCTGC 195
| | | | |
QY 61 CyPheSePProSeG1yH1a1eLeuH1aAaPArCySeSerThrAaPlyThrThrVal1e 80
| | | | |
Db 196 TGTTCCTCCCTTCAAGACATATTTGGCATGTGTGAACAAGATGTACCACTGTCTCTA 255
| | | | |
QY 81 TrPaAaThG1uAaG1y1eMeTLeuH1aVal1eMetG1uG1nProSeR1ySePProVal 100
| | | | |
Db 256 TGAATATACGAAATTAAGATGCTGGGCAAGATGGAACAGCTTAAGTGGCAGCCCTGTG 315
| | | | |
QY 101 ArGVal1yAaG1yPheSePProAaPArThrCySeLeuH1aSeG1yVal1a1aAaPlyThr 120
| | | | |
Db 316 AGGATTGCAAGTTTCCCAAGCTTCAAGTGTGGCATCAAGGCGCACTGATGAACT 375
| | | | |
QY 121 Val1a1eThrPaAaH1aG1nSerTyR1yLeuTyR1yArG1ySeR1ySeVal1yAaPly 140
| | | | |
Db 376 GTGGTTTGTGGAATGACACATCATCAATTAATATAGATGTGTGTTAAAGATGGC 435
| | | | |
QY 141 SeRLeuH1a1aCyAa1aPheSePProAaG1ySeRPhaPheVal1ThrG1ySeSerCyS 160
| | | | |
Db 436 TCCTTGCGGCAATGTCATTTCTCTTAATGGAAGCTTCTTGTCACTGCTCTCATGT 495
| | | | |


```
QY 161 GlyAspLeuThrValITrPAspAspLysMetArgCysLeuHisSerGluLysAlaHisAsp 180
    |||||
Db 496 GGTATTTTAAACAGTGGGATGATTAATATGATGGTGTCTGCATATGTAAGCAACATGAT 555
QY 181 LeuGlyIleThrCysCysAspPheSerSerGlnProValSerAspGlyValGlnGlyLeu 200
    |||||
Db 556 CTGGAAATTACCTGCTCGATTTTCTTCAACAGCCAGTTTCTGATGAGAAACAAGTCTT 615
QY 201 GlnPhePheArgLeuAlaSerCysGlyValAspCysGlnValLysIleTrpIleValSer 220
    |||||
Db 616 CAGTTTTCGACGTGGCATCATGTGTCAGGATTTGCCAAGTCAAAATTTGATTTGTTTCT 675
QY 221 PheThrIleIleLeu----- 225
    |||||
Db 676 TTTTACCATATCTTAGGTTTGAATTAATAATATAAGTACACTGAGTGGGCATGTGCT 735
QY 225 ----- 225
Db 736 CTTGTCTGGCTTGCTGCTTTTCCCATGATGGCAGATGCTAGTCAAGGTGATGAT 795
QY 225 ----- 225
Db 796 AAGTCTGTCAATGATATGATTAATACTGAGAAATATCTTCAACATTTGACTCAGCAC 855
QY 225 ----- 225
Db 856 ACCAGATATGCAACAATGCTTTTGACCTTAATACCTTTTACTTGTCTACTGTGTTCA 915
QY 226 -----AlaArgArg 228
    |||||
Db 916 ATGACAAATAACAGTGAACATCTGGCAATTGACCTGGAACAATTGGCAAGAGGCGC 975
QY 229 ThrGlnHisGlnLeuLysGlnPheThrGlnAspTrpSerGluGluValValSerThrTrp 248
    |||||
Db 976 ACAGAAATCATAGCTGAAGCAATTTTACCAAGATTTGTCAGAGGAGATGCTCTCAACATG 1035
QY 249 LeuCysAlaGlnAspLeuLysAspLeuValGlyIlePheLysMetAsnAsnIleAspGly 268
    |||||
Db 1036 CTTTGTGCACAAAGTTTAAAGATCTTGTGTGATTTTTCAGATGAAATTAACATTTGATGGA 1095
QY 269 LysGluLeuLeuLeuLeuThrLysGluSerLeuAlaAspPheLysIleGluSerLeu 288
    |||||
Db 1096 AAAGAACTGTGTAATCTTACAAAGAAAGTCTGCTGATGATTTGAAATTTGAATCTCTA 1155
QY 289 GlyLeuArgSerLysValLeuArgLysIleGluGluLeuArgThrLysValLysSerLeu 308
    |||||
Db 1156 GGACTGGGTAGTAAAGTCTGAGGAAATTAAGAGCTCAGAGACCAAGGTTAAATCCCTT 1215
QY 309 SerSerGlyIleProAspGluPheIleCysProIleThrArgGluLeuMetLysAspPro 328
    |||||
Db 1216 TCTTCAGGAATTCCTGATGAAATTTATATGTCATTAACCTAGAGAACTTATGAAAGATCCG 1275
QY 329 ValIleAlaSerAspGlyIleTrpSerLysGluLysGluAlaMetGluAsnTrpIleSerLys 348
    |||||
Db 1276 GTATCCCATCAGATGGCTATTCATATGAAAGAAAGAACATGAAATTTGGATCAGCAA 1335
QY 349 LysLysArgThrSerProMetThrAsnLeuValLeuProSerAlaValLeuThrProAsn 368
    |||||
Db 1336 AAGAAACATCAACTCCATGACCAAAATCTTGTCTTCCACGCGTATCTTACCAACAAAT 1395
QY 369 ArgThrLeuLysMetAlaIleAsnArgTrpLeuGlnThrHisGlnLys 384
    |||||
Db 1396 AGGACTCTGAAATAAGCCATCAATGATGCTGAGACACACCAAAAG 1443
```

RESULT 5
US-10-357-930-29046
; Sequence 29046, Application US/10357930
; Publication No. US20040255086A1

; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John

```
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF  
; FILE REFERENCE: MRI-007BCN  
; CURRENT APPLICATION NUMBER: US/10/357,930  
; PRIOR FILING DATE: 2003-02-04  
; PRIOR APPLICATION NUMBER: 09/785,276  
; PRIOR FILING DATE: 2003-02-16  
; PRIOR APPLICATION NUMBER: 60/183,319  
; PRIOR FILING DATE: 2000-02-17  
; PRIOR APPLICATION NUMBER: 60/189,862  
; PRIOR FILING DATE: 2000-03-16  
; PRIOR APPLICATION NUMBER: 60/207,454  
; PRIOR FILING DATE: 2000-05-25  
; PRIOR APPLICATION NUMBER: 60/211,314  
; PRIOR FILING DATE: 2000-06-09  
; PRIOR APPLICATION NUMBER: 60/219,007  
; PRIOR FILING DATE: 2000-07-18  
; PRIOR APPLICATION NUMBER: 60/255,281  
; NUMBER OF SEQ. ID NOS: 62232  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 29046  
; LENGTH: 1996  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc_feature  
; LOCATION: 1, 2, 3, 1987, 1988, 1989, 1990, 1991, 1992, 1993, 1994,  
; LOCATION: 1995, 1996  
; OTHER INFORMATION: n = A,T,C or G  
US-10-357-930-29046  
  
Alignment Scores:  
Pred. No.: 1,81e-225 Length: 1996  
Score: 1984.00 Matches: 383  
Percent Similarity: 80.46% Conservative: 0  
Best Local Similarity: 80.46% Mismatches: 1  
Query Match: 96.92% Indels: 92  
DB: 18 Gaps: 1  
  
US-10-077-111-13 (1-384) x US-10-357-930-29046 (1-1996)  
QY 1 MetValLysLeuLeuLeuThrLeuAlaAspHisGlyAspAspValaAspCysValaPhe 20  
    |||||  
Db 16 ATGTGAACTGATTCACATTAAGTGAATGATGATGATGATGATGATGATGATGATGATGATG 75  
QY 21 SerPheSerLeuLeuAlaThrCysSerLeuAspLysThrIleArgLeuIleTrpSerLeuArg 40  
    |||||  
Db 76 TCCCTTTCCCTCTGGCTACTTGTCTGCTTGGACAAACAAATTCGCTGATGCTGTTACGT 135  
QY 41 AspPheThrGluLeuProHisSerProLeuLysPheHisThrTyraIleValHisCysCys 60  
    |||||  
Db 136 GACTTTACTCAACTGACCACTTCTCCATTAAGTTTCATACATGCTGTCACCTGCTGC 195  
QY 61 CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrThrValLeu 80  
    |||||  
Db 196 TGTTCCTCCCTTCAGACATATTGTCGATCGGTTCACCAAGATGATGATGATGATGATGATG 255  
QY 81 TrpAsnThrGluAsnGlyGlnMetLeuAlaValMetGluGlnProSerGlySerProVal 100  
    |||||  
Db 256 TGAATATCTGAAATATGACAGATGCTGACGATGATGAAACAGCTTACGACGCTGCTGCTG 315  
QY 101 ArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyAlaAlaAspGlyThr 120  
    |||||  
Db 316 AGGTTTGCAGATTTTCCAGACTCCAGCTGTTGGCATCAGGGGCGAGCTGATGAACT 375  
QY 121 ValValLeuThrPheAlaGlnSerTyrlLysLeuTyrlArgCysGlySerValLysAspGly 140  
    |||||  
Db 376 GTGGTTTGTGATGACAGATCATTAATAATGATGATGATGATGATGATGATGATGATGATGATG 435  
QY 141 SerLeuAlaIleCysAlaPheSerProAspGlySerPhePheValThrGlySerSerCys 160  
    |||||
```

Db 436 TCCTGGCGGCGATGCGATTTTCTCCTAATGAGACTTCTTTGTCACTGGCTCCATGCT 495
Qy 161 GLYAspLeuThrValTTPAspApyMelaArgCysLeuHisSerGluValAlaHisAsp 180
Db 496 GGGGATTTAAACAGTGGTGGATGATTAATAAGGTGTCTGATAGTGAATAAGACATGAT 555
Qy 181 LeuGlyIleThrCysCysAspPheSerSerGlnProValSerApyGluGlnGlyLeu 200
Db 556 CTGGGAATTACCGCTGCGCATTTTCTTCCACAGCCGATTTCTGATGAGAACAGGCTCT 615
Qy 201 GlnPhePheArgLeuAlaSerCysGlyGlnApyCysGlnValIleTyrPheIleValSer 220
Db 616 CAGTTTTTTCAGCTGCGCATCATGTGGTCAGAGATTGGCAAGTCAAAATTTGGATTGTCTT 675
Qy 221 PheThrHisIleLeu----- 225
Db 676 TTTTACCATATCTTAGGTTTGAATTAATAATAAGTACACTGAGTGGCACTGTGCT 735
Qy 225 ----- 225
Db 736 CCTGTTCTGGCTTGTGCTTTTCCATGATGGCGAGATGCTAGTCTCAGGGTCAGTGAAT 795
Qy 225 ----- 225
Db 796 AAGTCTGTCAATGATATATGATTAATTAATCTGAGAAATTAATCTTCAACATTGACTCAGCAC 855
Qy 225 ----- 225
Db 856 ACCAGGTATGTCACAACTTGTGCTTTTGCACCTAATACCTTTTACTGTGCTAGTGTCA 915
Qy 226 -----AlaArg 228
Db 916 ATGACAAAACAGTGAACATCTGCAATTGACCTGGAAAACATTGCCCAAGCAAGCGCG 975
Qy 229 ThrGlnHisGlnLeuLeuValGlnPheThrGlnAspTyrSerGlnGluValIleSerThrTyr 248
Db 976 ACAGAACATCAAGTGAAGCAATTACCAGAAATTTGGTCAGAGAGAGATGTCTCAACATGG 1035
Qy 249 LeuCysAlaGlnAspLeuValAspLeuValGlyIlePheIleMetAsnAsnIleApyGly 268
Db 1036 CTTTGTGCAACAATTTAAAGATCTTGTGGTATTTTCAAGATGAATTAACATTTGATGCA 1095
Qy 269 LysGlnLeuLeuAsnLeuThrLysGlnSerLeuAlaAspAspLeuLysIleGlnSerLeu 288
Db 1096 AAAAGAACTGTTGAATCTTACAAAAGAAAGTGGCTGATGATTTGAAAATTTGAATCTCTA 1155
Qy 289 GlyLeuArgSerLysValLeuArgLysIleGlnGlnLeuArgThrLysValIleSerLeu 308
Db 1156 GGACTGCGTAGTAAGTGTGAGGAAATTTGAAGAGCTCAGACCAAGGTTAAATCCCTT 1215
Qy 309 SerSerGlyIleProAspGlnPheIleCysProIleThrArgGlnLeuMetLysAspPro 328
Db 1216 TCTTCAGAAATCTCTGATGAATTAATATGTCCAAATTAAGAACTTAAGAAAGATCCG 1275
Qy 329 ValIleAlaSerAspGlyTyrSerLysGlnLysGlnAlaMetGlnAsnTyrIleSerLys 348
Db 1276 GTCATGCGATCAGATGAGCTATTCATATGAAGAAAGAAAGCAATGAAATTTGATCAGAGAA 1335
Qy 349 LysLysArgThrSerProMetThrAsnLeuValLeuProSerAlaValLeuThrProAsn 368
Db 1336 AAGAAACGTACAGATCCCATGACAAATCTTGTCTTCCACGGTAATCTTACACCAAAAT 1395
Qy 369 ArgThrLeuLysMetAlaIleAsnArgTyrLeuGlnThrHisGlnLys 384
Db 1396 AGGACTCTGAAAATGGCCATCAATAGTGGCTGGAAGACACACCAAAAG 1443

RESULT 6
US-10-287-218-40
; Sequence 40, Application US//10287218
; Publication No. US20030198975A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: AZIMZAI, Yalda; AU-YOUNG, Janice K.

; APPLICANT: BATRA, Sajeev; BAUGHN, Mariah R.
; APPLICANT: BECHA, Shanya D.; BOROWSKY, Mark L.
; APPLICANT: BUFORD, Neill; DING, Li
; APPLICANT: ELLIOTT, Vicki S.; EMERLING, Brooke M.
; APPLICANT: GANDHI, Ameena R.; GIETZEN, Kimberly J.
; APPLICANT: GRIFFIN, Jennifer A.; HAFALIA, April J.A.
; APPLICANT: HONCHELU, Cynthia D.; LAL, Preeti G.
; APPLICANT: LEE, Soo Yeun; LU, Dyrng Alina M.
; APPLICANT: ARVIZU, Chandre S.; RAMKUMAR, Jayalaxmi
; APPLICANT: REDDY, Roopa; SANJANMALA, Madhu, M.
; APPLICANT: TANG, Y. Tom; WALIA, Narinder K.
; APPLICANT: WANG, Yu-mei, E.; WARREN, Bridget A.
; APPLICANT: XU, Yuming; YANG, Dunning
; APPLICANT: YAO, Monique G.; YUE, Henry
; APPLICANT: ZEBARJADIAN, Yeganeh
; TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH
; FILE REFERENCE: PI-0417 USA
; CURRENT APPLICATION NUMBER: US/10/287,218
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: PCT/US02/11152
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: US 60/349,705
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: US 60/295,263
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/295,340
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/293,727
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/291,846
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/291,662
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: US 60/287,228
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/286,820
; PRIOR FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: US 60/283,294
; PRIOR FILING DATE: 2001-04-11
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 40
; LENGTH: 1773
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030198975A1 7483131CB1
US-10-287-218-40
Alignment Scores:
Pred. No.: 7,8e-225 Length: 1773
Score: 1978.00 Matches: 382
Percent Similarity: 80.25% Conservative: 0
Best Local Similarity: 80.25% Mismatches: 2
Query Match: 96.63% Indels: 92
DB: 16 Gaps: 1
US-10-077-111-13 (1-384) x US-10-287-218-40 (1-1773)
Qy 1 MetValLysLeuIleHisThrLeuAlaAspHisGlyAspAspValAsnCysCysAlaPhe 20
Db 60 ATGGTAAACTAGTTACACATTAAGTATGATGATGATGATGATGATGATGATGATGATG 119
Qy 21 SerPheSerLeuLeuAlaThrCysSerLeuAspLysThrIleArgLeuTyrSerLeuArg 40
Db 120 TCCTTTCCCTCTTGGCTACTGCTCTTGGACAAACAAATTCGCTGTAATCGTTACGT 179
Qy 41 AspPheThrGlnLeuProHisSerProLeuLysPheHisThrTyrAlaValHisCysCys 60
Db 180 GACTTACTGAACTGCCACATTTCTCATTAAGTTCAATCACTAATGCTGCTGCTGCTGCTG 239

QY 61 CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrValLeu 80
DB 240 TGTTCCTCCCTCCAGACATATTTGGCATCGTGTTCACAGATGGTACCTGCTCTA 299
QY 81 ThrAspThrGluAsnGlyGlnMetLeuAlaValMetGluInProSerGlySerProVal 100
DB 300 TGGAAATACGAAATGACAGATGCTGGCAGTATGGAACAGCTTGTGGCACCCTGTG 359
QY 101 ArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyAlaAlaAspGlyThr 120
DB 360 AGGCTTGGCAGATTTTCCCACTCCAGACTGTTTGGCATCAGGGGAGCTGATGAGACT 419
QY 121 ValValLeuThrAsnAlaGlnSerTyrTyrLeuTyrArgCysGlySerValIleAspGly 140
DB 420 GTGGTTTGGATGCACAGTCATACAAATTAATATGATGTGTATGTTAAAGATGC 479
QY 141 SerLeuAlaAlaCysAlaPheSerProAsnGlySerPhePheValThrGlySerSerCys 160
DB 480 TCTTGGCGGCAATGATTTTCTCTTAATGAAGAGCTTCTTGTCACTGGCTCTCATGT 539
QY 161 GlyAspLeuThrValTTPAspAspLysMetArgCysLeuHisSerGlyLysAlaHisAsp 180
DB 540 GGTGATTTAACAGTGTGGATGATTAATAATAGGTGTCTGATGTGAAAAAGCAGCATGAT 599
QY 181 LeuGlyIleThrCysCysAspPheSerSerGlnProValSerAspGlyGlnGlyLeu 200
DB 600 CTGGATTAATCCTGCTGCGCATTTTCTTCAAGCAGATTTCTGATGGAACAAGGCTT 659
QY 201 GlnPhePheArgLeuAlaSerCysGlyGlnAspCysGlnValIleTyrIleValSer 220
DB 660 CAGTCTTTTGCAGTGGCATGATGTGTCAGAGATTTGCCAGATTCAAAATTTGATTTCTT 719
QY 221 PheThrHisIleLeu----- 225
DB 720 TTTAACCATATCTTAAGTTTGAATTAATAATAAGTCACTGAGTGGCAGCTGTGCT 779
QY 225 ----- 225
DB 780 CCGTTCGTGGCTTGTGCTTTTCCCATGATGGGAGATGCTAGTCTGAGGGTCACTGAT 839
QY 225 ----- 225
DB 840 AAGTCTGTCAATGATATGATACTAATACTAGAGATTAATCTGCACATTTGACTCAGAC 899
QY 225 ----- 225
DB 900 ACCAGTANGTCAACAATCTGTGCTTTGACCTTAATACCTTTTACTTGTCTACTGTCA 959
QY 226 -----AlaArgArg 228
DB 960 ATGACAAAACAGTGAACATCTGGCAATTTGACCTGGAACACTTTGCCAAGCAGAGAGC 1019
QY 229 ThrGlnHisGlnLeuLysGlnPheThrGlnAspTyrSerGlnGluValIleSerThrTyr 248
DB 1020 ACAGAACATCAGCTGAAGCAATTTTACCGAAGATTTGTCAGAGGAGATGTCTCAACATG 1079
QY 249 LeuCysAlaGlnPheLeuLysAspLeuValGlyIlePheLysMetLanAenIleAspGly 268
DB 1080 CTTTGTGCACAAATTTTAAAGATCTTGTGTATTTTCAAGATGAATTAACATTGATGGA 1139
QY 269 LysGlnLeuLeuAsnLeuThrLysGlnSerLeuAlaAspAspLeuLysIleGlnSerLeu 288
DB 1140 AAGAACTGTTGAATCTTACAAAGAAAGTCTGGCTATGATTTGAAAAATTGAATCTCTA 1199
QY 289 GlyLeuArgSerLysValLeuArgLysIleGlnGlnLeuArgThrLysValLysSerLeu 308
DB 1200 GGAAGTGGTGAATAAGTGAAGAAATGGAAGAGCTCAGAGACCAAGTTAAATCCCTT 1259
QY 309 SerSerGlyIleProAspGlnPheIleCysProIleThrArgGlnLeuMetLysAspPro 328
DB 1260 TCTTCAGAAATTCCTGTATGAAATTAATATGTCCTCAATACTGAGAACTTAAGAAAGATCCG 1319
QY 329 ValIleLeuAspArgGlyTyrSerTyrGlnLysGlnAlaMetGlnAenTyrIleSerLys 348

DB 1320 GTCATGCATCAGATGCTGATTTTCAATGAAGAAAGCAATGGAATTTGATCAGCAAA 1379
QY 349 LysLysArgThrSerProMetThrAsnLeuValLeuProSerAlaValLeuThrProAsn 368
DB 1380 AAGAAAGCTGACAGTCCCATGACAAATCTTGTCTTCTTCAAGCGGTACTTACACCAAT 1439
QY 369 ArgThrLeuLysMetAlaIleAsnArgTyrLeuGlnThrHisGlnLys 384
DB 1440 AGGACTCTGAAATGCGCATCATATGATGGCTGAGACACACAAAG 1487
RESULT 7
US-10-474-291-40
; Sequence 40, Application US/10474291
; Publication No. US20040132043A1
; GENERAL INFORMATION:
; APPLICANT: AZIMZAI, Yalda; AU-YOUNG, Janice K.
; APPLICANT: BATRA, Sajeev; BAUGHN, Mariah R.
; APPLICANT: BECHA, Shanya D.; BOROMSKY, Mark L.
; APPLICANT: BURFORD, Neil; DING, Li
; APPLICANT: ELLIOTT, Vicki S.; EMERLING, Brooke M.
; APPLICANT: GANDHI, Ameena R.; GIETZEN, Kimberly J.
; APPLICANT: GRIFFIN, Jennifer A.; HARALIA, April J.A.
; APPLICANT: HONCHELL, Cynthia D.; LAU, Preeti G.
; APPLICANT: LEE, Soo Yeun; LU, Dyung Aina M.
; APPLICANT: ARVIZU, Chandra S.; RAMKUMAR, Jayalaxmi
; APPLICANT: REDDY, Roopa M.; SANJUNWALA, Madhusudan M.
; APPLICANT: TANG, Y. Tom; CHAWLA, Narinder K.
; APPLICANT: WANG, Yu-Mei E.; WARREN, Bridget A.
; APPLICANT: XU, Yuming; YANG, Junning
; APPLICANT: YAO, Montague G.; YUE, Henry
; APPLICANT: ZEBARJADIAN, Veganeh
; TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH
; FILE REFERENCE: PI-0417 USN
; CURRENT APPLICATION NUMBER: US/10/474,291
; CURRENT FILING DATE: 2003-10-06
; PRIOR APPLICATION NUMBER: PCT/US02/11152
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: US 60/349,705
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: US 60/295,263
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/295,340
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/293,727
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/291,846
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/291,662
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: US 60/287,228
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/286,820
; PRIOR FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: US 60/283,294
; PRIOR FILING DATE: 2001-04-11
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 40
; LENGTH: 1773
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 7483131CB1
US-10-474-291-40
Alignment Scores:
Pred. No.: 7.8e-225 Length: 1773
Score: 1978.00 Matches: 382
Percent Similarity: 80.25% Conservative: 0
Best Local Similarity: 80.25% Mismatches: 2

Query Match: 96.63% Indels: 92
DB: 18 Gaps: 1
US-10-077-111-13 (1-384) x US-10-474-291-40 (1-1773)

QY 1 MetValIyLeuIleHISThrLeuAlaAerHISGlyAAspAspValIenCyCybAlaPhe 20
DB 60 ATGGTAAACGTGATTCACATTAGCATGAGCATGATGCAACGTGTCCTTC 119
QY 21 SerPheSerLeuLeuAlaThrCySerSerLeuAspLyThrIleArgLeuYrSerLeuArg 40
DB 120 TCTTTTCCCTCTTGCTACTTCTCTCTTGACAAAACAATTGGCTGTACTGTACGT 179
QY 41 AspPheThrGluLeuProHISerProLeuLyPheHISThrTyAlaValHISCyCyb 60
DB 180 GACTTACTGAACGTGCACATCTCCATTGAACTTCACTACATGCTGCTGCACCTGCTG 229
QY 61 CybPheSerProSerGlyHISIleLeuAlaSerCySerThrAspGlyThrThrValLeu 80
DB 240 TGTTCCTCCCTTCAGAGACATATTTGGCATCGTGTTCACAAGATGGTACACTGTCTTA 299
QY 81 TrpAsnThrGluLeuGlyGlnMetLeuAlaValMetGlnGlnProSerGlySerProVal 100
DB 300 TGGAAATCTGAAATATGACAGATGCTGGCATGAGACAGCTTGTGGCAGCCCTGTG 359
QY 101 ArgValICyGlnPheSerProAspSerThrCybLeuAlaSerGlyAlaIaAspGlyThr 120
DB 360 AGGGTTTGGCAGTTTCCCCAGACTCCAGCTGTTGGCATCAAGGGGAGCTGATGGAACT 419
QY 121 ValIleuLeuTrpAsnAlaGlnSerTyLybLeuTyArgCybGlySerValIyAspGly 140
DB 420 GTGGTTTGTGGAAATGACAGCTCATCAAAATTAATAGATGTGTATTAAGATGGC 479
QY 141 SerLeuAlaIaCybAlaPheSerProAsnGlySerPheValIThrGlySerSerCyb 160
DB 480 TCTTTGGCGCATGTGATTTCTCTTAATGAAAGCTTTCTTGTACTGCTCTCATGT 539
QY 161 GlyAspLeuThrValItrpAspAspLyMetArgCybLeuHISerGlyLybAlaHISAsp 180
DB 540 GGTGATTTAAACAGTGTGGATGATTAATAATGAGTGTCTGATGTGAAAAAGCATGAT 599
QY 181 LeuGlyIleThrCybCybAspPheSerSerGlnProValSerAspGlyGlnGlnLeu 200
DB 600 CTTGGAAATTAACCTGCTGCGCATTTTCTTCACAGCCAGTTTCTGATGAGAAACAAGGCTT 659
QY 201 GlnPhePheArgLeuAlaSerCybGlyGlnAspCybGlnValIySerIleTrpIleValSer 220
DB 660 CAGTTTTCGACTGGCATCATGTGATGAGATTTGCCAAGTCAAAATTTGGATTGTCT 719
QY 221 PheThrHISIleLeu----- 225
DB 720 TTTAACCATATCTTTAGGTTTGAATTAATAATATAAGTACAGTAGGGGACATGTGCT 779
QY 225 ----- 225
DB 780 CCTGTCTGGCTTGCTTTTCCATGATGGCAGATGCTAGTCTCAGGGTCAATGAT 839
QY 225 ----- 225
DB 840 AAGTCTGTCAATAGTATATGATTAATACTAGAGAAATATACTTACACATTGACTCAGCAC 899
QY 225 ----- 225
DB 900 ACCAGTATGTCAACAATTGTGCTTTGACATTAACCTTTTACTTGTACTGTGCTCA 959
QY 226 -----AlaArgArg 228
DB 960 ATGGAACAAAACAGTGAACATCTGGCAATTGACCTGGAAACATTTGGCAAGAGAGC 1019
QY 229 ThrGlnHISGlnLeuLybGlnPheThrGluAspTrpSerGlyGlnValIaSerThrTrp 248
DB 1020 ACAGAAACATCAAGCTGAAGCAATTTTACCGAAGATTGTGCAAGAGAGATGTCTCAACATGG 1079

QY 249 LeuCybAlaGlnAspLeuLybAspLeuValGlyIlePheLybMetAsnAsnIleAspGly 268
DB 1080 CTTTGTGCAACAAGATTAAAGATCTTGTGGTATTTTCAGATGAATTAACATTTGATGGA 1139
QY 269 LybGluLeuLeuAsnLeuThrLybGlySerLeuAlaAspAspLeuYrIleGlySerLeu 288
DB 1140 AAGAACTTTGAATTTTCAAAAGAAAGTCTGGCTGATGATTTGAAAAATTGAATCTCTA 1199
QY 289 GlyLeuArgSerLybValLeuArgLybIleGlnGluLeuArgThrLybValIySerLeu 308
DB 1200 GAGCTGCTAGTAAAGTCTGAGAAATTTGAAGAGCTCAGGACCAAGGTTAAATCCCTT 1259
QY 309 SerSerGlyIleProAspGluPheIleCybProIleThrArgGluLeuMetLybAspPro 328
DB 1260 TCTTCAGGAATTCCTGATGAATTTATATCTCCAAATTAAGAACTTATGAAAGATCCG 1319
QY 329 ValIleAlaSerAspGlyTySerTyArgLybGluAlaMetGlnAsnTrpIleSerLyb 348
DB 1320 GTCATGCAATCAGATGGCTATTCATATGAAAGAAAGCAATGGAATTTGATCACCAAA 1379
QY 349 LybLybArgThrSerProMetThrAsnLeuValLeuProSerAlaValLeuThrProAsn 368
DB 1380 AAGAAACGTACAAGTCCCATGACAAATCTTGTCTTCTTCAGCGGTACTTACACCAAA 1439
QY 369 ArgThrLeuLybMetAlaIleAsnArgTrpLeuGluThrHISGlnLyb 384
DB 1440 AGGACTGTGAAATGGCCATCAATAGATGCTGAGAGACACACCAAAAG 1487

RESULT 8

US-10-104-047-1689
; Sequence 1689, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1689
; LENGTH: 1811
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-1689

Alignment Scores:

Pred. No.: 8,06e-225 Length: 1811
Score: 1978.00 Matches: 382
Percent Similarity: 80.25% Conservative: 0
Best Local Similarity: 80.25% Mismatches: 2
Query Match: 96.63% Indels: 92
DB: 17 Gaps: 1

US-10-077-111-13 (1-384) x US-10-104-047-1689 (1-1811)

QY 1 MetValIyLeuIleHISThrLeuAlaAerHISGlyAAspAspValIenCyCybAlaPhe 20
DB 150 ATGGTAAACGTGATTCACATTAGCATGAGCATGATGCAACGTGTCCTTC 209
QY 21 SerPheSerLeuLeuAlaThrCySerSerLeuAspLyThrIleArgLeuYrSerLeuArg 40
DB 210 TCTTTTCCCTCTTGCTACTTCTCTCTTGACAAAACAATTGGCTGTACTGTACGT 269
QY 41 AspPheThrGluLeuProHISerProLeuLyPheHISThrTyAlaValHISCyCyb 60
DB 270 GACTTACTGAACGTGCACATCTCCATTGAACTTCACTACATGCTGCTGCACCTGCTG 329
QY 61 CybPheSerProSerGlyHISIleLeuAlaSerCySerThrAspGlyThrThrValLeu 80
DB 330 TGTTCCTCCCTTCAGAGACATATTTGGCATCGTGTTCACAAGATGGTACACTGTCTTA 389

Qy 81 TrpAnthrGluAnGlyGlnMetLeuAlaValMetGluGlnProSerGlySerProVal 100
Db 390 TGGAAATCTGAAAATGACAGATGCTGGCACTGATGAAACGCTTAAGTGGACGCCCTGTG 449
Qy 101 ArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyAlaIAspGlyThr 120
Db 450 AGGGTTGGCAGTTTCCCGACACTCCAGTGTGGATCAGGGGGCAGCTGATGGAAC 509
Qy 121 ValValLeuThrPanaAlaGlnSerThrLysLeuYrArgCysGlySerValIAspGly 140
Db 510 GTGGTTTGTGGAATGACAGCTCATACAAATTAATAGATGTGGATGTTAAAGATGGC 569
Qy 141 SerLeuAlaIaCysAlaPheSerProAspGlySerPhePheValThrGlySerSerCys 160
Db 570 TCTCTGGGGCAGATGCTATTTCTCTTATGAAAGCTTTCTTGACCTGGCTCTCATGT 629
Qy 161 GlyAspLeuThrValITrpAspAspLysMetArgCysLeuAlaSerGlyLysAlaIAsp 180
Db 630 GGTGATTTAACAGGTGGGATGATAAATGAGGTGTCTGATGAAAAAGCACATGAT 689
Qy 181 LeuGlyIleThrCysGAspPheSerSerGlnProValSerAspGlyGluGlnGlyLeu 200
Db 690 CTTGGAATTACCTCTGCTGGATTTTCTTACAGCCAGTTTCTGATGGAGAACAGGCTTT 749
Qy 201 GlnPhePheArgLeuAlaSerCysGlyGlnAspCysGlnValLysIleThrIleValSer 220
Db 750 CAGTTTTCAGCTGGCATCATGTGTGAGATGTGCCAAGTTCAAATTTGGATTTGTTCT 809
Qy 221 PheThrIleIleLeu----- 225
Db 810 TTTACCATATCTTAGTTTGAATTAATAATAAAGTACAGTACAGTGGCAGCTGTCT 869
Qy 225 ----- 225
Db 870 CCTGTTCGCTGTGTCTTTTCCCATGATGGCAGATGTAAGTCTCAGGGTCACTGGAT 929
Qy 225 ----- 225
Db 930 AAGTCTGTCATGATATGATATGATACTGAGAAATATATCTTCAACATGACTGACAGC 989
Qy 225 ----- 225
Db 990 ACCAGTATGTCACAACTTGTGCTTTGACCTTAATACCTTTTACTGTACTGTCTCA 1049
Qy 226 -----AlaArgArg 228
Db 1050 ATGGAACAAACAGTGAACATCTGGCAATTGACCTGAAACACTTGGCCAAAGCAGC 1109
Qy 229 ThrGluHISGlnLeuLysGlnPheThrGluAspTrpSerGluGluValIAspThrTrp 248
Db 1110 ACAGAACATCAGCTGAGAACAAATTAACGAAATGTGTGAGAGAGATGTCTCAACATGG 1169
Qy 249 LeuCysAlaGlnAspLeuLysAspLeuValGlyIlePheLysMetAsnAsnIleAspGly 268
Db 1170 CTTTGTGACAAAGATTTTAAAGATCTTTGTGTATTTTCAAGATGAATAACATTTGATGCA 1229
Qy 269 LysGluLeuLeuAsnLeuThrLysGluSerLeuAlaAspAspLeuLysIleGluSerLeu 288
Db 1230 AAAGAACCTTGAATCTTACAAAGAAAGTCTGGCTGATGATTTGAAATTTGAATCTCTTA 1289
Qy 289 GlyLeuAspSerLysValLeuArgLysIleGluGluLeuAspGlnThrLysValLysSerLeu 308
Db 1290 GGACTGCTAGTAAAGTCTGAGGAAATGAGAGCTCAGAGACCAAGTTAAATCCCTT 1349
Qy 309 SerSerGlyIleProAspGluPheIleCysProIleThrArgGluLeuMetLysAspPro 328
Db 1350 TCTTCAGAAATTCCTGATGATTTTATATGTCCAATTAAGAACTTATGAAAGATCCG 1409
Qy 329 ValIleAlaSerAspGlyTrpSerTrpGluLysGluAlaMetGluAsnTrpIleSerLys 348
Db 1410 GTCAATCGCATCAGATGGCTATTCAATATGAAAGAAAGCAATGAAATTTGGATGACGAA 1469

Qy 349 LysLysArgThrSerProMetThrAsnLeuValLeuProSerAlaValLeuThrProAsn 368
Db 1470 AAGAAAGTCAAGATCCCATGACAAATCTTGTCTTCTTCCACAGGATTAACCAAAAT 1529
Qy 369 ArgThrLeuLysMetAlaIleAsnArgTrpLeuGluThrIleGlnLys 384
Db 1530 AGGACTCTGAAAATGGCCATCAATAGATGGCTGGAGACACCAAAAG 1577

RESULT 9
US-09-971-392-207
; Sequence 207, Application US/09971392
; Publication No. US20030134283A1
; GENERAL INFORMATION:
; APPLICANT: Peterson, David P.
; APPLICANT: Pearson, Cecelia I.
; APPLICANT: Cocks, Benjamin G.
; TITLE OF INVENTION: GENES REGULATED IN DENDRITIC CELL DIFFERENTIATION
; FILE REFERENCE: PA-0029 US
; CURRENT APPLICATION NUMBER: US/09/971,392
; CURRENT FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: 60/237,652
; PRIOR FILING DATE: 2000-10-03
; NUMBER OF SEQ ID NOS: 260
; SOFTWARE: PERL Program
; SEQ ID NO 207
; LENGTH: 1817
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Template ID: 158923.9
US-09-971-392-207

Alignment Scores:
Pred. No.: 8,11e-225 Length: 1817
Score: 1978.00 Matches: 382
Percent Similarity: 80.25% Conservative: 0
Best local Similarity: 80.25% Mismatches: 2
Query Match: 96.63% Indels: 92
DB: Gaps: 1

US-10-077-111-13 (1-384) x US-09-971-392-207 (1-1817)

Qy 1 MetValLysLeuIleHISThrLeuAlaAspHISGlyAspAspValaAsnCysAlaPhe 20
Db 148 ATGGTGAACCTGATTCACACATTAAGTATGATGATGATGATGATGATGATGATGATG 207
Qy 21 SerPheSerLeuLeuAlaThrCysSerLeuAspLysThrIleArgLeuYrSerLeuArg 40
Db 208 TCTTTTCCCTCTTGGCTACTGTCTCTTGGACAAACATTCGCTGTACTGTACTGTACT 267
Qy 41 AspPheThrGluLeuProHISerProLeuLysPheHISThrYrAlaValHISCysCys 60
Db 268 GACTTTCAGTAACTGACACATTTCTCATTAAGTTTCAATGATGATGATGATGATGATG 327
Qy 61 CysPheSerProSerGlyHISIleLeuAlaSerCysSerThrAspGlyThrThrValLeu 80
Db 328 TGTTCCTCCCTTCAGACATATTTTGGCATCGGTTCACAGATGGTATCCACTGTCTTA 387
Qy 81 TrpAnthrGluAnGlyGlnMetLeuAlaValMetGluGlnProSerGlySerProVal 100
Db 388 TGGAAATCTGAAAATGACAGATGCTGGCACTGATGAAACGCTTAAGTGGACGCCCTGTG 447
Qy 101 ArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyAlaIAspGlyThr 120
Db 448 AGGGTTGGCAGTTTCCCGACACTCCAGTGTGGATCAGGGGGCAGCTGATGGAAC 507
Qy 121 ValValLeuThrPanaAlaGlnSerThrLysLeuYrArgCysGlySerValIAspGly 140
Db 508 GTGGTTTGTGGAATGACAGCTCATACAAATTAATAGATGTGGATGTTAAAGATGGC 567
Qy 141 SerLeuAlaIaCysAlaPheSerProAspGlySerPhePheValThrGlySerSerCys 160

```
Db      568 TCCTTGCGCGGATGTCATTTTCTCCTAATGAAAGCTTCTTGTCACTGGCTCTCANTG 627
Qy      161 G1yAspLeuThrVal1TpaAspAllyMeArGcYseuH1sSerG1uYsa1aH1sAsp 180
Db      628 GGTGATTTAAcAGTGGATGATTAATAAGGTCGTGATATGTGAAAAAGCAcATGAT 667
Qy      181 LeuG1y1LeThrcYsAspPheSerSerGlnProValSerAspG1yG1uG1nG1yLeu 200
Db      688 CTTCGAAATTACCTGCGCCGATTTTTCTTTCACACCCAGATTCTGATGGAGAACAGGCTT 747
Qy      201 GlnPhePheArgLeuAlaSerCySG1yGlnAspCySerGlnVal1yS1eTPr1LeVal1Ser 220
Db      748 CAGTTTTTTCAGCTGGCATATGTGTCAGAGATTGCCAAGTCAAAATTTCGATTGTTTCT 807
Qy      221 PheThH1s1LeLeu----- 225
Db      808 TTAAcCATATCTTAAAGTTTGAATTAAATATTAAGTACACGTAGTGGGCACTGTGCT 867
Qy      225 ----- 225
Db      868 CCGTTTCTGGCTTGTGCTTTTCCATGATGGGCAAGTGTAGTCTCAGGGTCAGTGAAT 927
Qy      225 ----- 225
Db      928 AAGTCTGTCACTAGTATATGATTAATACTAGAGAAATTAATTCTTACACATTTGACTCAGCAC 967
Qy      225 ----- 225
Db      988 ACCAGATGTGCACAACCTGTGCTTTTGACACTAATACCTTTTAACTTGCTACTGATGTTCA 1047
Qy      226 -----11AATGATG 228
Db      1048 ATGAGCAAAAACAAGTGAACATCTGGCAATTGACCTGGAAAACATTTGCCAAGCAAGAGGC 1107
Qy      229 ThcG1uH1sG1nLeuYsG1nPhethrGluAspTrpSerG1uG1uVal1Val1SerThrTrp 248
Db      1108 ACGAAACATCACTGAGACAAATTACCGAAGATTTGTCAGAGAGAGATGTCCTCAACATGG 1167
Qy      249 LeuCyAlaGlnAspLeuYsAspLeuVal1G1y1LePheYsMeArAsnEn1LeAspG1y 268
Db      1168 CTTTGTGCAACAAGATTTAAAGATCTTGTGATATTTTCAAGATGAAATACATTTGATGGA 1227
Qy      269 LysG1uLeuLeuAsnLeuThr1yLysG1uSerLeu1aAspAspLeuYs11eG1uSerLeu 288
Db      1228 AAAGAACCTGTGAATCTTCAAAAAGAAAGCTGCTGATATTTGAAAATTGAATCTCTTA 1287
Qy      289 G1yLeuArGSerYsVal1LeuArgYs11eG1uG1uLeuArGThr1ySVal1ySserLeu 308
Db      1288 GGACTGCTAGTAAGAGCTGAGAGAAATTTGAAGAGCTCAGACCAAGTTAAATCCCTT 1347
Qy      309 SerSerG1y1LeProAspG1uPhe1LeCySPro11eThrArgG1uLeuMeT1ySAspPro 328
Db      1348 TCTTCAGGAATTCCTGATGAATTTATATATGTCCAATACCTAGAGAACTTATGAAAGATCCG 1407
Qy      329 Val11eAlaSerAspG1yTrpSerTrpG1uLysG1uAlaMeT1uAsnTrp11eSerYs 348
Db      1408 GTCATGCAACAATGCTATTCATATGAAAAGAAAGCAATGGAATAATTCAGATCAGCAAA 1467
Qy      349 LysLysArGThrSerProMetThrAsnLeuVal1LeuProSerAlaVal1LeuThrProAsn 368
Db      1468 AAGAAACGTACAAAGTCCCATGACAAATCTTGTTCCTTCACGGGACTTACCAcCAAT 1527
Qy      369 ArgThrLeuYsMeAr1a1eAsnArgTrpLeuG1uThrH1sG1nYs 384
Db      1528 AGGACCTGAAAATGGCCATCAATAGATGCTGAGAGCAcACCAAAAG 1575

RESULT 10
US-10-037-270-768
; Sequence 768, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
```

```
; APPLICANT: Aeundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunding
; APPLICANT: Wang, Dunru
; APPLICANT: Wang, Zhwei
; APPLICANT: Tillinghast, John
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/10/037,270
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pc_FL_genes Version 1.0
; SEQ ID NO 768
; LENGTH: 1844
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: CDS
; LOCATION: (164)..(1594)
US-10-037-270-768

Alignment Scores:
Pred. No.: 8.3e-225 Length: 1844
Score: 1978.00 Matches: 382
Percent Similarity: 80.25% Conservative: 0
Best Local Similarity: 80.25% Mismatches: 2
Query Match: 96.63% Indels: 92
DB: 15 Gaps: 1

US-10-077-111-13 (1-384) x US-10-037-270-768 (1-1844)
Qy      1 MeetVal1yLeu11eH1sThrLeu1aAspH1sG1yAspAspVal1AsnCySAlaPhe 20
Db      164 ATGTGAACCTGATTTTACACATTTAGCTGATCATGTGACGATGTCACTGCTGCTTC 223
Qy      21 SerPheSerLeuLeu1aThrcYsSerLeuAsp1yThr1LeArgLeuYsSerLeuArg 40
Db      224 TCTTTTCCCTCTTGCTGCTGCTGCTGCTTGGACAAACAAATTCGCTGACTGTTAGCT 283
Qy      41 AspPheThrG1uLeuProH1sSerProLeuYsPheH1sThrTrpAlaVal1H1sCyS 60
Db      284 GACTTATCTGAACCTGCAcATTCCTCAATTGAAGTTTCAACTATGCTGTGCTCACTGCTGC 343
Qy      61 CySPheserProSerG1yH1s1LeLeu1aSerCySserThrAspG1yThrThrVal1Leu 80
Db      344 TGTTCCTCCCTTCAGGACATATTTTGGCATGCTGTTCAACAGATGTTCCACTGCTCTTA 403
Qy      81 TrpAsnThrG1uAsnG1yG1uMeT1eLeu1aVal1MeT1uG1nProSerG1ySerProVal 100
Db      404 TGAATATCTGAAAATGGAACAGATGCTGGCAGATGAGAAACAACCTAGTGGAGCCCTGG 463
Qy      101 ArgVal1CySG1nPheserProAspSerThrcYsLeu1aSerG1yAla1aAspG1yTr 120
Db      464 AGGATTTCAGATTTTCCCAcAGCTCCAGCTGTTGGGCATCAGGGGCAAGCTGATGAACT 523
Qy      121 Val1Val1LeuThrAsnAlaG1nSerTrpYsLeuYTrArgCySG1ySerVal1ySAspG1y 140
Db      524 GTGTTTGTGGAATGCAAGTCATCAAAATTAATATATAGTGTGTGATGTTAAAGATGGC 583
```

QY	141	serLeuAlaAlaCyAlaPheSerProAsnGlySerPhePheValThrGlySerSerCyS	160
Db	584	TCCTTGGCGGAGATGACATTTCTCTTAATGAAACCTCTTTGTCACTGGCTCTCACTG	643
QY	161	GlyAspLeuThrValTTPAspAspLysMetArgCySLeuHisSerGlyLysAlaHisAsp	180
Db	644	GGTGATTTAAcAGTGTGGGATGATAAATAGAGTGTCTGCATATGTAAGAAAAGCACATAT	703
QY	181	LeuGlyIleThrCySAspPheSerSerGlnProValSerAspGlyGluGlnGlyLeu	200
Db	704	CTTGGAATTAACCTGCTCGATTTCTTCTTCAAGCAGATTTCTGATGAGGAACAAGTCTT	763
QY	201	GlnPhePheAspGlyLeuAlaSerCyGlyGlnAspCySglnValLysIleLeuProIleValSer	220
Db	764	CAGTTTTTTTCAGCTGGCAGTCATATGTTGTCAGAAATGCCAAATCAAAATTTGAAATGTTTCT	823
QY	221	PheThrHisIleLeu-----	225
Db	824	TTTAAcCATATCTTAAGTCTTTGAATTAATAATTAATAAGTACACAGTGAAGGCACGTGCT	883
QY	225	-----	225
Db	884	CCTGTTCTGCTGCTGCTTTTCCCATGATGGCAGATGTAGTCTCAGAGGTCAATGGAT	943
QY	225	-----	225
Db	944	AAAGTGTGATAGTATATGATATCTAATATCTGAGAAATATATCTTCAACATATGACCTGACAC	1003
QY	225	-----	225
Db	1004	ACCAAGTATGTCACAACTTGTCCTTTGACACCTAATACCCTTTTACTTGCTACTGTTCA	1063
QY	226	-----AlaArgArg-----	228
Db	1064	ATGACACAACAGTGAACATCTGGCAATTTGACCTGCAACACTTTGGCAAGCAAGAGAC	1123
QY	229	ThrGluHisGlnLeuLysGlnPheThrGluAspTrpSerGluGluValIleSerThrTrp	248
Db	1124	ACAAACATCTACGCTGAAAGCAATTTAACTCCAAAGATTTGGTCAAGAGAGATGTCTCAACATGG	1183
QY	249	LeuCyAlaGlnAspLeuLysAspLeuValGlyIlePheLysMetAsnAsnIleAspGly	268
Db	1184	CTTTGTGCACAAAGATTTAAAGATTTTGTGGTATTTCAAGATTAATACATTCATGCA	1243
QY	269	LysGluLeuLeuAsnLeuThrLysGluSerLeuAlaAspAspLeuLysIleGluSerLeu	288
Db	1244	AAAGACCTGTTGAATCTTACAAAGAAAGTCTGCGTATGATTTGAAAATTGAATCTCTA	1303
QY	289	GlyLeuArgSerLysValLeuArgLysIleGluGluLeuValGlnLysValLysSerLeu	308
Db	1304	GGACTGCGCTAATAAGTCTGAGGAAATTAAGAGAGCTCAGAGCCAAAGTTAAATCCCTT	1363
QY	309	SerSerGlyIleProAspGluPheIleCySProIleThrArgGluLeuMetLysAspPro	328
Db	1364	TCTTCAGGAATTCCTGATGAATTAATATGTCCAATTAAGTAAAGAACTTAATGAAGATCCG	1423
QY	329	ValIleAlaSerAspGlyTrpSerTyrgLysGluAlaMetGluAsnTrpIleSerLys	348
Db	1424	GTCAATCGATCAGATGCGTATTTCAATATGAAAAGAAAGAACATGAAATTTGATCTGCAAA	1483
QY	349	LysLysArgThrSerProMetThrAsnLeuValLeuProSerAlaValLeuThrProAsn	368
Db	1484	AAAGAAACGTAAAGCCCATGACAAATCTTGTCTTCTTCAAGCGGTACTTAACCAAT	1543
QY	369	ArgThrLeuLysMetAlaIleAsnArgTrpLeuGluThrHisGlnLys	384
Db	1544	AGGACTCTGAAAATGGCCATCATATGATGGCTGAGGACACACAAAG	1591
RESULT 11			
US-10-117-722-768			
; Sequence 768, Application US/10117722			
; Publication No. US20030219744A1			
; GENERAL INFORMATION:			

```

APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. US2003021974A1el Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 784C1P2BC1P
CURRENT APPLICATION NUMBER: US/10/117,722
CURRENT FILING DATE: 2002-04-04
PRIOR APPLICATION NUMBER: 09/620,312
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1104
SOFTWARE: pc_Fl_genes Version 1.0
SEQ ID NO 768
LENGTH: 1844
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (164)..(1594)
US-10-117-722-768

```

Alignment Scores:	
Pred. No.:	8.3e-225
Length:	1
Score:	1978.00
Matches:	0
Percent Similarity:	80.25%
Conservative:	0
Best Local Similarity:	80.25%
Mismatches:	2
Query Match:	96.63%
Indels:	1
DB:	17
Gaps:	1

US-10-077-111-13 (1-384) X US-10-117-722-768 (1-1844)

Qy	1	MeValIysLeu11e1isThrIeu1a1sPh1sG1yAaPaPVal1Asn1CyS1Al1Phe	20
Dd	164	ATGGTGAACCTGATTCACACATTACTGATCATGGTACAGATGCACACTGGTGGCTTC	22
Qy	21	SePheSer1eu1a1aThr1CyS1Ser1eu1sThr11e1a1g1eu1Y1Ser1eu1a1g	40
Dd	224	TCTTTTCCCTCTGGCTACTCTTGCTCTTGACAAACAAATTCGCTGTACTCTGTACGT	28
Qy	41	AspPheThrG1u1eu1Pro1isSer1Pro1eu1y1Phe11sThr1Y1a1a1a1H1s1CyS1s	60
Dd	284	GACTTTACTGAACCTGCCACATTCCTCATTTGAAAGTTTCATTACTTATGCTGTCACTGCTGC	34
Qy	61	CyAPheSer1ProSer1YH1s1le1eu1a1aSer1CyS1Ser1h1ra1sG1Y1Thr1Val1eu	80
Dd	344	TGTTTCTCCCTTCAGACATATTTTGGCATGTGTTCAACAGATGGTACACACTGTCTTA	40
Qy	81	ThrAsnThrG1u1a1nG1yG1Met1eu1a1a1a1MetG1u1n1ProSerG1y1Ser1Pro1a1	100
Dd	404	TGGAAATACGAAAAATGACAGATGCTGGCAGATGAGGAACCGCTAAGGCGAGCCCTGTG	46
Qy	101	ArgVal1CyS1G1n1Phe1Ser1Pro1Asp1Ser1Thr1CyS1eu1a1a1a1a1a1sP1G1Y1Thr	120
Dd	464	AGGGTTTGGCAGATTTTCCCCAGACTCCACGATGTTGGCATCAGGGGGCAGCTGATGAAC	52
Qy	121	Val1Val1eu1ThrAsn1a1a1n1Ser1Y1r1y1s1eu1Y1a1a1G1y1Ser1Val1y1a1sP1G1Y	140
Dd	524	GAGGTTTTTGGAATGACAGATCAACAAATTAATACATGTGTGAAGTTAAAGAAAGTGC	58
Qy	141	Se1Leu1a1a1CyS1a1a1Phe1Ser1Pro1Asn1y1Ser1Phe1Phe1Val1n1G1y1Ser1Ser1CyS	160
Dd	584	TCTTTGGCGGAGTGCATTTTCTCTAATGAAGACTCTTTGTCACTGCTCTCTCAGT	64
Qy	161	GlyAsP1eu1Thr1Val1TrpAsp1sP1y1s1e1a1a1G1y1s1e1u1H1s1SerG1u1y1a1a1H1a1sP	180
Dd	644	GGGATTTTAACTGTGGATGATTAATTAAGAGTCTGCATAGTGAAGAAAGACATGAT	70

Db 856 CCTGTTGGCTTGTGCTTTTCCCATGATGGCAGATGTAAGTCTCAGGGTCAGTGAT 915
QY 225 ----- 225
Db 916 AAGTCTGCATAGTATATGACTAATACTGAAATATCTTCACACATTGACTCAGCAC 975
QY 225 ----- 225
Db 976 ACCAGATATGTCACAACTTGCTTTGGCACCCTAATACCCTTTACTTGCTACTGTTCA 1035
QY 226 ----- 226
Db 1036 ATGACAAACAGTGAACATCTGGCAATTTGACCTGAAACATTTGGCAAGGAGGCGC 1095
QY 229 ThrluhlslnleuLysGlnPheThrGluAspTrpSerGluGluValIleSerThrTrp 248
Db 1096 ACAGAACCTCAGCTGAGACCAATTTACCAAGATTGGTATGAGAGAGATGCTTCACACATGG 1155
QY 249 LeuCysAlaGlnAspLeuLysAspLeuValGlyIlePheLysMetAsnAsnIleAspGly 268
Db 1156 CTTGTGCAACAAGATTTAAAGATCTTGTAATTTTCAAGATGAATTAACATTGATGGA 1215
QY 269 LysGluLeuLeuAsnLeuThrLysGluSerLeuAlaAspAspLeuLysIle----- 285
Db 1216 AAGAAGCTGTGAATCTTACAAAGAAAGCTGGCTGATGATTTGAAATTTGGCTGGAGT 1275
QY 285 ----- 285
Db 1276 CCTTGGCAGTGTATGCTTCATGCGACCTTCAACCTCTGGGCTCAAGTATCCTCCTA 1335
QY 286 -----GluSerLeuGlyLeuArgSerLysValLeuArgLysIleGluGluLeuArg 302
Db 1336 CCTGGGCTCATCTCTAGAGCTGCGTAAAGTCTGAGAGAAATTTGAAGACTCAGG 1395
QY 303 ThrLysValLysSerLeuSerSerGlyIleProAspGluPheLysCysProIleThrArg 322
Db 1396 ACCAAGGTGAATCCCTTCTTCAGGAATCTGATGAATTTATATGCTCAATTAATCAGA 1455
QY 323 GluLeuMetLysAspProValIleAlaSerAspGlyTrpSerGlyGluLysGluAlaMet 342
Db 1456 GAACCTTGAAGAAGATCCGGTATGCGACATGCGATGCTATTAATGAAGAAAGCAATG 1515
QY 343 GluAsnTrpLysSerLysLysLysArgThrSerProMetThrAsnLeuValLeuProSer 362
Db 1516 GAAATTTGATCAGCAAAAAGAAAGCTACAGTCCCATGACAAATCTTTGCTCTTCA 1575
QY 363 AlaValLeuThrProAsnArgThrLeuLysMetAlaIleAsnArgTrpLeuGluThrHis 382
Db 1576 GCGGTACTTACACCAATAGAGACTGTGAATGGCCATCATATGAGCTGGAGACACAC 1635
QY 383 GlnLys 384
Db 1636 CAAAG 1641

RESULT 13
US-10-077-111-6
; Sequence 6, Application US/10077111
; Publication No. US20020187492A1
; GENERAL INFORMATION:
; APPLICANT: Todderud, C. Gordon
; APPLICANT: Finger, Joshua N.
; APPLICANT: Rillema, Jili
; TITLE OF INVENTION: TBA
; FILE REFERENCE: 3053-4114US2
; CURRENT APPLICATION NUMBER: US/10/077,111
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 60/294,181
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: 60/269,366
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6

LENGTH: 1901
; TYPE: DNA
; ORGANISM: MOUSE
US-10-077-111-6
Alignment Scores:
Pred. No.: 1,07e-181 Length: 1901
Score: 1616.50 Matches: 312
Percent Similarity: 72.06% Conservative: 31
Best Local Similarity: 65.55% Mismatches: 40
Query Match: 78.97% Indels: 93
DB: 13 Gaps: 3
US-10-077-111-13 (1-384) x US-10-077-111-6 (1-1901)
QY 1 MetValLysLeuIleHisThrLeuAlaAspHisGlyAspAspValaenyCysAlaPhe 20
Db 19 ATGTGAGGTGATTCACAGCTGTGATGATCAGGGAGATGACCTCAGCTCTGCGCTTC 78
QY 21 SerPheSerLeuLeuAlaThrCysSerLeuAspLysThrIleArgLeuTrpSerLeuArg 40
Db 79 TCGGCTGCCCTCTGCGCACCTGCTCTTGGACAAAGCAATCCGTCTGTACTCCCTAAGT 138
QY 41 AspPheThrGluLeuProHisSerProLeuLysPheHisThrTyralaValHisCysCys 60
Db 139 GACTTGTGAAGTCCGCTACTCCCGCTGAAGTTCACACCTATGCTGTCACTGCTGC 198
QY 61 CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrValLeu 80
Db 199 TGTTCACCCCTCAGACACAGTTTATGATCTGTGTCAGACAGCGGACACCGTCTG 258
QY 81 TrpAsnThrGluAsnGlyGlnMetLeuAlaValMetGluGlnProSerGlySerProVal 100
Db 259 TGGAGCTGCAACGCGACACACCTGACCGTGTGGAGAGCCGGGTGGAGCCCTGTG 318
QY 101 ArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyAlaAlaAspGlyThr 120
Db 319 CGGCTGTGCTTTTCCAGACTGCTGCTACTACGTCAGGGGCTGCGATGATCC 378
QY 121 ValValLeuTrpAsnAlaGlnSerTyrlsLeuTyrlsArgCysGlySerValLysAspGly 140
Db 379 ATTGCTTTGGGAATGACACAGACATCAAACTATATGAGTGTATGTCAAGGAATGC 438
QY 141 SerLeuAlaAlaCysAlaPheSerProAsnGlySerPheValThrGlySerSerCys 160
Db 439 TCATGTGCTGCTGTGCTTTTCCCGATGAGGCTCTTTGCTACTGCTCTCGGCGC 498
QY 161 GlyAspLeuThrValTrpAspAspLysMetArgCysLeuHisSerGlyLysAlaHisAsp 180
Db 499 GGGGACTTGAACAGTGGATGACAGATAGAGTGTATACACAGCGAAGGCGCACGAT 558
QY 181 LeuGlyIleThrCysCysAspPheSerSerGlnProValSerAspGlyGluGlnGlyLeu 200
Db 559 CTGGGATCAGCTGCTCAGCTTTTCTTCTCAGAGCTCTCTGCGGAGAA---GGCCTC 615
QY 201 GlnPheAspArgLeuAlaSerCysGlyGlnAspCysGlnValLysIleTrpIleValSer 220
Db 616 CAGCTTACCAAGTTGGGTGATGTGTCAGAGCTGTGAATCAAACTGCGGCTTACT 675
QY 221 PheThr----- 222
Db 676 ATTAACCGCTGTAGGCTTTGAATTAATATTAAGACACACTAAGTGGGCACTGGCC 735
QY 222 ----- 222
Db 736 CCTGTTGCGCTGTGCTTTTTCATGATGAAAGATGCTTGATCGGGGTAGTGAT 795
QY 223 -----HisIleLeuAlaArgArg 228
Db 796 AAATCTGTCAATACATGATGATCGCCCTCAGAGTGTGCTACACAGCTGACTCAGCAT 855
QY 229 ThrGluHis----- 231

Db 856 ACCAGTATGTTACGACTTGGCTTGGACCAACACTCTTACTTGGTTCA 915
QY 231 ----- 231
Db 916 ATGACAGACAGTGAACATTGGAGCTTGACCTGGAAACACCTTGCAGAGAGAGC 975
QY 232 -----GlnLeuLysGlnPheThrGluAspTrpSerGluGluValIleSerThrTrp 248
Db 976 ATGACAGACCCGGTGAACATTTTCACATGAAATGGACAGAGAGATGCTCTCCGTGG 1035
QY 249 LeuCyAlaIleGlnAspLeuLysAspLeuValGlyIlePheLysMetAsnAsnIleAspGly 268
Db 1036 CTTTCGTGCTTAAGGCTTGGAAAGCTTCGTGATTTTTCAGGCAACACATCATGAGG 1095
QY 269 LysGluLeuLeuAsnLeuThrLysGluSerLeuAlaAspAspLeuLysIleGluSerLeu 288
Db 1096 AAGAACTATTGCAATCTCAAAAGAAAGCTGGCTGGTATTTGAAAATCGAATCTCTA 1155
QY 289 GlyLeuAspSerLysValLeuArgLysIleGluGluLeuArgThrLysValLysSerLeu 308
Db 1156 GGGCTGCGCAGCAAGCTCGAGAGATTTGAAGAGCTCAGGGCCAAAGATGATTCCTC 1215
QY 309 SerSerGlyIlePheProAspGluPheIleCysProIleThrArgGluLeuMetLysAspPro 328
Db 1216 TCTTCGGAAATCCCTGACGATTCATCTGCCAATAACAGAGAACTCATGAAGAGACCC 1275
QY 329 ValIleAlaSerAspGlyTrpSerTrpGluLysGluLysMetGluAsnTrpIleSerLys 348
Db 1276 GTCATGCAATCAAGTGGCTACTCTCTACGAGAGAAAGCAATGAAAGCTGATCCACAG 1335
QY 349 LysLysArgThrSerProMetThrAsnLeuValLeuProSerAlaValLeuThrProAsn 368
Db 1336 AAGAAAGGTACGAGCCCATGACAAATTTGGCTCTCCCTTCACTGTACTGACCCCAAC 1395
QY 369 ArgThrLeuLysMetAlaIleAsnArgTrpLeuGlnThrIleGlnLys 384
Db 1396 AGGACCTGAAGATGGCCATCAACGATGGCTGGAGCGACAGAGAG 1443

RESULT 14
US-10-077-111-5
; Sequence 5, Application US/10077111
; Publication No. US20020187492A1
; GENERAL INFORMATION:
; APPLICANT: Toddard, C. Gordon
; APPLICANT: Finger, Joshua N.
; APPLICANT: Rillema, Jill
; TITLE OF INVENTION: TBA
; FILE REFERENCE: 3053-4114US2
; CURRENT APPLICATION NUMBER: US/10/077,111
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 60/294,181
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: 60/269,366
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 630
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; OTHER INFORMATION: 630 bp partial nucleic acid sequence of human
; US-10-077-111-5

Alignment Scores:
Pred. No.: 2,78e-74 Length: 630
Score: 709.00 Matches: 148
Percent Similarity: 64.17% Conservative: 15
Best Local Similarity: 58.27% Mismatches: 41
Query Match: 34.64% Indels: 50
DB: 13 Gaps: 2

US-10-077-111-13 (1-384) x US-10-077-111-5 (1-630)
QY 98 SerProValArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyAlaAla 117
Db 19 GCTCCGCTTCTGGCTTGTGCTTTTCCATGATGGGCAATGCTAGTCTCAGGGTCAAG 78
QY 118 AspGlyThrValValLeuThrPheAsnIleGlnSerTrpLysLeuTrpArgCysGlySerVal 137
Db 79 GATTAATCTGTCATGATATATGATATCTAATTAATCTGAGAAATATCTTCACATGATCTCAG 138
QY 138 LysAspGlySerLeuAlaIleCysAlaPheSerProAsnGlySerPhePheValThrGly 157
Db 139 CACACAGGTATGTCAACCTTGCTTTTGGACCTTAATACCTTTTACTTGGTACTGAT 198
QY 158 SerSerCysGlyAspLeuThrValTrpAspAspLysMetArgCysLeuHisSerGluLys 177
Db 199 TCAATGACAAAACAGTGAACATCTGGCAATTTGACCTGGAAACATT----- 246
QY 178 AlaHisAspLeuGlyIleThrCysAspPheSerSerGlnProValSerAspGlyGlu 197
Db 246 ----- 246
QY 198 GlnGlyLeuGlnPhePheArgLeuAlaSerCysGlyGlnAspCysGlnValLysIleTrp 217
Db 247 -----TGGCA----- 252
QY 218 IleValSerPheThrHisIleLeuAlaArgArgTrpGlnIleGlnLeuLysGlnPheThr 237
Db 253 -----GCAAGGCCACAGAAACATCACTGAAAGCAATTTACC 288
QY 238 GluAspTrpSerGluGluValValSerThrTrpLeuCysAlaIleGlnAspLeuLysAspLeu 257
Db 289 GAAAGTTGGTCAAGAGAGATGCTCAACATGGCTTTGGCACAAGATTTAAAGATCTT 348
QY 258 ValGlyIlePheLysMetAsnAsnIleAspGlyLysGluLeuLeuAsnLeuThrLysGlu 277
Db 349 GTTGTATTTTCAAGATGATTAACATTTGATGAAAGAACTGTGAAATTTCAAAAGAA 408
QY 278 SerLeuAlaAspAspLeuLysIleGluSerLeuGlyLeuArgSerLysValLeuArgLys 297
Db 409 AGTCTGGCTGATGATTTGAAAATTTGAAATCTTAAGACTCGTAGTAAGTCTGAGAGAA 468
QY 298 IleGluGluLeuArgThrLysValLysSerLeuSerSerGlyIleProAspGluPheIle 317
Db 469 ATTGAAGACTCAGAGCAAGTTAATCCCTTTCTTCAGGAATCTTGATGAATTTAA 528
QY 318 CysProIleThrArgGluLeuMetLysAspProValIleAlaSerAspGlyTrpSerTrp 337
Db 529 TGTCCAAATTAACTAAGAACTTAATGAAGATCCGGTCACTGCATGCATGCTATTCATAT 588
QY 338 GluLysGluAlaMetGluAsnTrpIleSerLysLysValArg 351
Db 589 GAAAGAGAGCAATGGAATTTGATCAAGAAAAGAAACGT 630

RESULT 15
US-10-357-930-5029/c
; Sequence 5029, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862

; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 5029
; LENGTH: 446
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-5029

Alignment Scores:
Pred. No.: 8,33e-65 Length: 446
Score: 627.50 Matches: 122
Percent Similarity: 96.06% Conservative: 0
Best Local Similarity: 96.06% Mismatches: 2
Query Match: 30.65% Indels: 3
DB: 18 Gaps: 1

US-10-077-111-13 (1-384) x US-10-357-930-5029 (1-446)

QY 102 ValCySGln-PheSerProAspSer--ThrCysLeuAlaSerGlyAlaAlaAspGlyTh 120
|||
Db 444 GTTGGCCAGTTTCCCGACGCTCCCAACGTTTGGCATCAGGGCAGCTGATGAAAC 385

QY 120 rValValLeuTrpAsnAlaGlnSerTyrlsLeuTyrlsGlySerValIysAspGly 140
|||
Db 384 TGtGATTTTGGAATGCAAGTCATACAAATTATATGATGTGTAGTGTAAAGATGG 325

QY 140 ySerLeuAlaAlaCysAlaPheSerProAsnGlySerPheValThrGlySerSerCy 160
|||
Db 324 CTCCTGGCGGCATGTGATTTTCTCCATAGGAAGCTTTTGTCACTGCTCTCATG 265

QY 160 sGlyAspLeuThrValTrpAspAspLysMetArgCysLeuHisSerGlu-LysAlaHisA 180
|||
Db 264 TGGTGAATTTAACAGTGGGATGATAAATGAGGTGTCTGCATAGTGAATAAAGCACATG 205

QY 180 sPLeuGlyIleThrCysAspAspPheSerSerGlnProValSerAspGlyGlnGlnGlyL 200
|||
Db 204 ATCTTGAATTAACCTGCTGCGATTTTCTTCAACAGCCAGTTTCTGATGAGAACAGGCTC 145

QY 200 eugInPhePheArgLeuAlaSerCysGlyGlnAspCysGlnValIysIleTrpIleValS 220
|||
Db 144 TTCAGTTTTCGACTGGCATCATGTGTCTCAGATTGCCAAGTCAAAATTTGGATTGTTT 85

QY 220 erPheThrHisIleLeu 225
|||
Db 84 CTTTACCATATCTTA 68

Search completed: February 5, 2005, 15:08:28
Job time : 719 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 5, 2005, 08:35:31 ; Search time 748 Seconds
(without alignments)
9787.839 Million cell updates/sec

Title: US-10-077-111-12

Perfect score: 1272
Sequence: 1 gaattcggttcctacctcg.....acaccaaagtaagaattc 1272

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4313806 seqs, 2877871033 residues

Total number of hits satisfying chosen parameters: 3483234

Minimum DB seq length: 0

Maximum DB seq length: 200

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US10F_NEW_PUB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	176.8	13.9	180	9	US-09-864-761-27687, A Sequence 27687, A
C 2	132	10.4	132	9	US-09-864-761-30542 Sequence 30542, A
C 3	87	6.8	92	15	US-10-106-698-3381 Sequence 3381, App
C 4	36	2.8	36	13	US-10-077-111-19 Sequence 19, App
C 5	33.6	2.6	152	18	US-10-425-115-14085 Sequence 14085, A
C 6	31	2.4	165	17	US-10-242-535A-49505 Sequence 49505, A
C 7	31	2.4	165	17	US-10-085-783A-49505 Sequence 49505, A
C 8	30.6	2.4	136	9	US-09-770-696-357 Sequence 357, App
C 9	29.6	2.3	136	9	US-09-864-761-27694 Sequence 27694, A
C 10	29.2	2.3	177	18	US-10-437-963-57611 Sequence 57611, A
C 11	29	2.3	170	10	US-09-960-706-190 Sequence 190, App

12	29	2.3	180	15	US-10-238-075-979 Sequence 979, App
13	28.8	2.3	152	9	US-09-864-761-21108 Sequence 21108, A
14	28.8	2.3	153	9	US-09-864-761-19055 Sequence 19055, A
C 15	28.6	2.2	182	17	US-10-242-535A-34626 Sequence 34626, A
C 16	28.6	2.2	182	17	US-10-085-783A-34626 Sequence 34626, A
C 17	28.4	2.2	60	10	US-09-908-978-12618 Sequence 12618, A
C 18	28.2	2.2	128	17	US-10-424-599-64092 Sequence 64092, A
C 19	28	2.2	128	13	US-10-027-632-76001 Sequence 76001, A
C 20	28	2.2	143	17	US-10-027-632-76001 Sequence 76001, A
C 21	28	2.2	128	13	US-10-027-632-76003 Sequence 76003, A
C 22	28	2.2	143	17	US-10-027-632-76003 Sequence 76003, A
C 23	28	2.2	147	10	US-09-814-353-5092 Sequence 5092, App
C 24	28	2.2	147	10	US-09-814-353-11384 Sequence 11384, App
C 25	28	2.2	174	18	US-10-674-124A-66801 Sequence 66801, App
C 26	28	2.2	174	18	US-10-425-115-23798 Sequence 23798, A
C 27	28	2.2	186	18	US-10-437-963-83715 Sequence 83715, A
C 28	28	2.2	197	10	US-09-468-147-151 Sequence 151, App
C 29	28	2.2	197	17	US-10-319-745-151 Sequence 151, App
C 30	27.8	2.2	170	9	US-09-815-242-2740 Sequence 2740, App
C 31	27.8	2.2	170	9	US-09-815-242-3443 Sequence 3443, App
C 32	27.8	2.2	170	17	US-10-282-122A-5261 Sequence 5261, App
C 33	27.8	2.2	170	17	US-10-282-122A-6027 Sequence 6027, App
C 34	27.8	2.2	183	17	US-10-433-140A-1 Sequence 1, Appli
C 35	27.8	2.2	183	17	US-10-433-140A-3 Sequence 3, Appli
C 36	27.6	2.2	174	18	US-10-425-115-31475 Sequence 31475, A
C 37	27.6	2.2	184	9	US-09-864-761-20195 Sequence 20195, A
C 38	27.4	2.2	146	18	US-10-437-963-92321 Sequence 92321, A
C 39	27.4	2.2	192	9	US-09-796-692-83322 Sequence 8322, App
C 40	27.4	2.2	192	14	US-10-040-862-83322 Sequence 8322, App
C 41	27.4	2.2	192	17	US-10-057-475B-83322 Sequence 8322, App
C 42	27.4	2.2	192	17	US-10-154-884B-83322 Sequence 8322, App
C 43	27.4	2.2	182	18	US-10-764-324-83322 Sequence 8322, App
C 44	27.4	2.2	195	17	US-10-242-535A-2852 Sequence 2852, App
C 45	27.4	2.2	195	17	US-10-085-783A-2852 Sequence 2852, App

ALIGNMENTS

RESULT 1
US-09-864-761-27687/C
; Sequence 27687, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wenshang
; TITLE OR INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OR INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30

```
/ PRIOR APPLICATION NUMBER: PCT/US01/00668
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00663
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00662
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00661
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00670
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: US 60/234,687
/ PRIOR FILING DATE: 2000-09-21
/ PRIOR APPLICATION NUMBER: US 09/608,408
/ PRIOR FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: US 09/774,203
/ PRIOR FILING DATE: 2001-01-29
/ NUMBER OF SEQ ID NOS: 49117
/ SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
/ SEQ ID NO 27687
/ LENGTH: 180
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO AC006501.5
/ OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.88
/ OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
/ OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.81
/ OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.9
/ OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.73
/ OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.72
/ OTHER INFORMATION: EST_HUMAN HIT: AL040518.1, EVALUE 1.00e-94
/ OTHER INFORMATION: NT HIT: U43139.1, EVALUE 1.10e-01
/ OTHER INFORMATION: SWISSPROT HIT: Q16760, EVALUE 6.80e-02
US-09-864-761-27687
```

```
Query Match      13.9%; Score 176.8; DB 9; Length 180;
Best Local Similarity 98.9%; Pred. No. 5.8e-41;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 787 CAAGGCGCACAGACATCTGACGAGCAATTTACCGAAGTTGCTCAGAGAGATGCTCT 846
DB 180 CAAGGCGCACAGACATCTGACGAGCAATTTACCGAAGTTGCTCAGAGAGAGATGCTCT 121
QY 847 CAACATGGCTTTGTGCAACAAGATTAAAGATCTTTGTGTATTTTCAAGATGAATPACA 906
DB 120 CAACATGGCTTTGTGCAACAAGATTAAAGATCTTTGTGTATTTTCAAGATGAATPACA 61
QY 907 TTGATGAAAAGAACTGTTGAATCTTACAAAAGAAAGTCTGCTGATGATTGAAAATTG 966
DB 60 TTGATGAAAAGAACTGTTGAATCTTACAAAAGAAAGTCTGCTGATGATTGAAAATTG 1
```

```
RESULT 2
US-09-864-761-30542/c
/ Sequence 30542, Application US/09864761
/ Patent No. US20020048763A1
/ GENERAL INFORMATION:
/ APPLICANT: Penn, Sharon G.
/ APPLICANT: Rank, David R.
/ APPLICANT: Hanzel, David K.
/ APPLICANT: Chen, Wensheng
/ TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
/ FILE REFERENCE: Aecm1ca-X-1
/ CURRENT APPLICATION NUMBER: US/09/864,761
/ CURRENT FILING DATE: 2001-05-23
/ PRIOR APPLICATION NUMBER: US 60/180,312
/ PRIOR FILING DATE: 2000-02-04
/ PRIOR APPLICATION NUMBER: US 60/207,456
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: US 09/632,366
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: GB 24263.6
```

```
/ PRIOR FILING DATE: 2000-10-04
/ PRIOR APPLICATION NUMBER: US 60/236,359
/ PRIOR FILING DATE: 2000-09-27
/ PRIOR APPLICATION NUMBER: PCT/US01/00666
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00667
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00664
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00669
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00665
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00668
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00663
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00662
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00661
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00670
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: US 60/234,687
/ PRIOR FILING DATE: 2000-09-21
/ PRIOR APPLICATION NUMBER: US 09/608,408
/ PRIOR FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: US 09/774,203
/ PRIOR FILING DATE: 2001-01-29
/ NUMBER OF SEQ ID NOS: 49117
/ SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
/ SEQ ID NO 30542
/ LENGTH: 132
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO AC009307.1
/ OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.76
/ OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.73
/ OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.79
/ OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.7
/ OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.78
/ OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.89
/ OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.75
/ OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.68
/ OTHER INFORMATION: EST_HUMAN HIT: BE675766.1, EVALUE 2.00e-68
/ OTHER INFORMATION: NT HIT: X71133.1, EVALUE 1.20e+00
/ OTHER INFORMATION: SWISSPROT HIT: P54860, EVALUE 1.00e+00
US-09-864-761-30542
```

```
Query Match      10.4%; Score 132; DB 9; Length 132;
Best Local Similarity 100.0%; Pred. No. 6.1e-28;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 976 GACTGCGTAGTAAGTCTGAGAAAATGGAAGAGCTCAGACCAAGTTAATCCCTTT 1035
DB 132 GACTGCGTAGTAAGTCTGAGAAAATGGAAGAGCTCAGACCAAGTTAATCCCTTT 73
QY 1036 CTTGAGGAATTCCTGATGAATTTATATGTCCAAATACTNAGAACTTATGAAGAATCCG 1095
DB 72 CTTGAGGAATTCCTGATGAATTTATATGTCCAAATACTNAGAACTTATGAAGAATCCG 13
QY 1096 TCATGCAATCAG 1107
DB 12 TCATGCAATCAG 1
```

```
RESULT 3
US-10-106-698-3381
/ Sequence 3381, Application US/10106698
/ Publication No. US20030109690A1
/ GENERAL INFORMATION:
/ APPLICANT: Ruben et al.
```

;; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
;; FILE REFERENCE: PA005P1
;; CURRENT APPLICATION NUMBER: US/10/106,698
;; CURRENT FILING DATE: 2002-03-27
;; PRIOR APPLICATION NUMBER: PCT/US00/26524
;; PRIOR FILING DATE: 2000-09-28
;; PRIOR APPLICATION NUMBER: US 60/157,137
;; PRIOR FILING DATE: 1999-09-29
;; PRIOR APPLICATION NUMBER: US 60/163,280
;; PRIOR FILING DATE: 1999-11-03
;; NUMBER OF SEQ ID NOS: 8564
;; SOFTWARE: PatentIn Ver. 3.0
;; SEQ ID NO 3381
;; LENGTH: 92
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (45)..(45)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: misc_feature
;; LOCATION: (78)..(78)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: misc_feature
;; LOCATION: (86)..(86)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: misc_feature
;; LOCATION: (89)..(89)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: misc_feature
;; LOCATION: (91)..(91)
;; OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-3381

Query Match
Best Local Similarity 6.8%; Score 87; DB 15; Length 92;
Best Local Similarity 94.6%; Pred. No. 7.1e-15;
Matches 87; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 591 GGTGATTACAGCTGGGATGATTAATGAGTGTCTGCATAGTAAAGCAGCAT 650
DB 1 GGTGATTACAGCTGGGATGATTAATGAGTGTCTGCATAGTAAAGCAGCAT 60

QY 651 CTGGAATTACCTGCTGCGATTTTCTTCA 682
DB 61 CTGGAATTACCTGCTGCGATTTTCTTCA 92

RESULT 4
US-10-077-111-19/c
;; Sequence 19, Application US/10077111
;; Publication No. US20020187492A1
;; GENERAL INFORMATION:
;; APPLICANT: Toddard, C. Gordon
;; APPLICANT: Finger, Joshua N.
;; APPLICANT: Rillema, Jill
;; TITLE OF INVENTION: TBA
;; FILE REFERENCE: 3053-4114US2
;; CURRENT APPLICATION NUMBER: US/10/077,111
;; CURRENT FILING DATE: 2002-02-15
;; PRIOR APPLICATION NUMBER: 60/294,181
;; PRIOR FILING DATE: 2001-05-29
;; PRIOR APPLICATION NUMBER: 60/269,366
;; PRIOR FILING DATE: 2001-02-16
;; NUMBER OF SEQ ID NOS: 25
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 19
;; LENGTH: 36
;; TYPE: DNA
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: Primer, JNF
US-10-077-111-19

Query Match
Best Local Similarity 2.8%; Score 36; DB 13; Length 36;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1230 GCCATCAATAGATGCTGGAGACACCAAAAGTAA 1265
DB 36 GCCATCAATAGATGCTGGAGACACCAAAAGTAA 1

RESULT 5
US-10-425-115-140885/c
;; Sequence 140885, Application US/10425115
;; Publication No. US20040214272A1
;; GENERAL INFORMATION:
;; APPLICANT: La Rosa, Thomas J.
;; APPLICANT: Kovacic, David K.
;; APPLICANT: Zhou, Yihua
;; APPLICANT: Cao, Yongwei
;; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;; FILE REFERENCE: 38-21(53222)B
;; CURRENT APPLICATION NUMBER: US/10/425,115
;; CURRENT FILING DATE: 2003-04-28
;; NUMBER OF SEQ ID NOS: 369326
;; SEQ ID NO 140885
;; LENGTH: 152
;; TYPE: DNA
;; ORGANISM: Zea mays
;; FEATURE:
;; OTHER INFORMATION: Clone ID: MRT4577_5996C.1
US-10-425-115-140885

Query Match
Best Local Similarity 2.6%; Score 33.6; DB 18; Length 152;
Best Local Similarity 56.2%; Pred. No. 39;
Matches 63; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 323 GTGTTCAACAGATGATGATCACTGCTCATGGAATPACTGAAATGACAGATGTCAGT 382
DB 148 GTTTCACACAGAAATGATGATCACTGATGATGAAGCAGATTCCTGATACAGATGT 89

QY 383 GATGGAACAGCCTAGTGGCAGCCCTGTGAGGCTTTCAGTTTCCCGAC 434
DB 88 TATGCATCACCGAGCGGCTTGAACGGGAGGATTCGACAGTGGCCCGACAGC 37

RESULT 6
US-10-242-535A-49505
;; Sequence 49505, Application US/10242535A
;; Publication No. US20040013663A1
;; GENERAL INFORMATION:
;; APPLICANT: ChondroGene Inc.
;; APPLICANT: Liew, C.C.
;; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
;; FILE REFERENCE: 4231/2005
;; CURRENT APPLICATION NUMBER: US/10/242,535A
;; CURRENT FILING DATE: 2002-09-12
;; PRIOR APPLICATION NUMBER: US 10/085,783
;; PRIOR FILING DATE: 2002-02-28
;; PRIOR APPLICATION NUMBER: US 60/305,340
;; PRIOR FILING DATE: 2001-07-13
;; PRIOR APPLICATION NUMBER: US 60/275,017
;; PRIOR FILING DATE: 2001-03-12
;; PRIOR APPLICATION NUMBER: US 60/271,955
;; PRIOR FILING DATE: 2001-02-28
;; NUMBER OF SEQ ID NOS: 58994
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 49505
;; LENGTH: 165
;; TYPE: DNA
;; ORGANISM: Human
US-10-242-535A-49505


```
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 27694
; LENGTH: 136
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005228.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.86
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.84
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.75
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.9
; OTHER INFORMATION: NT HIT: M22967.1, EVALUO 1.20e+00
; OTHER INFORMATION: EST_HUMAN HIT: BF352940.1, EVALUO 2.00e-18
; US-09-864-761-27694

Query Match      2.3%; Score 29.6; DB 9; Length 136;
Best Local Similarity 53.4%; Pred. No. 5.4e+02;
Matches 62; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY      876 GATCTGTGTTGATTTTCAAGATGATTAACATTTGATGAGAAAGAACTGTTGAATCTTACA 935
DB      17 GTTGTATTATGTACCTTTCACATGAGAGAAATTCCTGGAAGAAATTTGTTGATCATGGA 76

QY      936 AAGAAAGCTGCTGCTGATGATTGAAATTTGAATCTCTAGAGACTGCGTAGTAAGT 991
DB      77 AGAAACATTTGTTATCATGCTTCAGAAATGAGATTGAGATGCGCTGCAATTAAT 132

RESULT 10
US-10-437-963-57611
; Sequence 57611, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovacic, David K.
; APPLICANT: Zhou, Yinhua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barabak, Brad
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; PRIOR FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 57611
; LENGTH: 177
; TYPE: DNA
; ORGANISM: Oryza sativa
```

```
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_59405C.1
; US-10-437-963-57611

Query Match      2.3%; Score 29.2; DB 18; Length 177;
Best Local Similarity 57.8%; Pred. No. 8.4e+02;
Matches 52; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY      757 TTGGATGTTGTTCTTTTACCCATATCTTACAGAGCGCACAGAAATCAGCTGAAGCAAT 816
DB      63 TTGGAAATATTTATGTTTACAGAGTTTATCCGCAAGCCAGACAGATATATCATTTGACATTT 122

QY      817 TTACCGAAGATTGGTCAGAGAGGTCGCTCT 846
DB      123 TCACCTGAGAGTATTCAGAGGGTATCTAT 152

RESULT 11
US-09-960-706-190/c
; Sequence 190, Application US/09960706
; Publication No. US20030134280A1
; GENERAL INFORMATION:
; APPLICANT: Munger, William E.
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplasia
; FILE REFERENCE: 44921-5029-01US
; CURRENT APPLICATION NUMBER: US/09/960,706
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 60/223,323
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 09/873,319
; PRIOR FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 190
; LENGTH: 170
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20030134280A1 AA279028
; US-09-960-706-190

Query Match      2.3%; Score 29; DB 10; Length 170;
Best Local Similarity 55.4%; Pred. No. 9.3e+02;
Matches 56; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY      863 ACAAGATTTAAAGATCTGTGTTATTTTCAAGATGATTAACATTTGATGAGAAAGAACT 922
DB      121 ACAATTTCTAGTAATATTTCTGAGTAAATTTTCTGATGATTAATTTCTAGATGATTAAGTTCT 62

QY      923 GTTGAATCTTACAAAGAAAGCTGCTGATGATTTGAAA 963
DB      61 GTTTAAAGCGCAATTTATTTTGAAGTTGTAACTTAATAA 21

RESULT 12
US-10-238-075-979
; Sequence 979, Application US/10238075
; Publication No. US20030148324A1
; GENERAL INFORMATION:
; APPLICANT: I. N. S. E. R. M.
; TITLE OF INVENTION: Polynucleotides which are of nature B2/D+ A- and which are isolated
; FILE REFERENCE: BLANDINE
; CURRENT APPLICATION NUMBER: US/10/238,075
; CURRENT FILING DATE: 2002-09-10
; PRIOR APPLICATION NUMBER: 0003145
; PRIOR FILING DATE: 2000-03-10
; NUMBER OF SEQ ID NOS: 1576
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 979
; LENGTH: 180
; TYPE: DNA
```

```
; ORGANISM: Escherichia coli
US-10-238-075-979

Query Match      2.3%; Score 29; DB 15; Length 180;
Best Local Similarity 49.7%; Pred. No. 9.7e+02;
Matches 74; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 557 TAATGAAAGCTTTTGTCACTGCTCCATCATGTGTGATTAACAGTGTGGATGATAA 616
    |||||
DB 21 TATTTAAACCGGAATGCGCGGTGAGATTTTCGAGATGAATTCGAAGTGTGATGATGATGC 80
    |||||

QY 617 AATGAGGTGTCTGATAGTAAAGACATGATCTTGAATTAACCTGCTGCATTTTTC 676
    |||||
DB 81 GAGCGCTGATGAAATACATACAGCAAGCAAAAGATTTTCTTAACACTACGCAATTAATC 140
    |||||

QY 677 TTCACAGCCAGTTTCTGATGAGAACAG 705
    |||||
DB 141 ATATCAAGAAATTAAGACGAAAGAGAG 169

RESULT 13
US-09-864-761-21108
; Sequence 21108, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 21108
```

```
; LENGTH: 152
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007721.10
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.95
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.96
; OTHER INFORMATION: EST HUMAN HIT: BE142985.1, EVALUATE 2.00e-80
; OTHER INFORMATION: NT HIT: AB012193.1, EVALUATE 1.00e-80
; OTHER INFORMATION: SWISSPROT HIT: Q13619, EVALUATE 3.00e-23
US-09-864-761-21108

Query Match      2.3%; Score 28.8; DB 9; Length 152;
Best Local Similarity 50.0%; Pred. No. 1e+03;
Matches 72; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 777 CATATCTTAGCAAGCGCCACAGAACTCATGCTGAAGCAATTTACGGAATTTGTAATTTTCAG 836
    |||||
DB 6 CAGTCCCTGCGCTGTGGCAAGCACGTGTGATTTAAAGTCCCAAGAAAGAAAGTG 65
    |||||

QY 837 GAGTGTCTCAACATGCGCTTTGTGCACAAGATTTAAAGATCTTGTGTAATTTTCAG 896
    |||||
DB 66 GAAGATGAGAGACAGATTCAATTTTAATGAGAGTTCAAGCAAGATTGTTTGAATTAAG 125
    |||||

QY 897 ATGAATACATTGATGAAAGAA 920
    |||||
DB 126 ATCAATCAATTTCAGATGAAGAA 149
    |||||

RESULT 14
US-09-864-761-19055
; Sequence 19055, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
```

PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 19055
LENGTH: 153
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC007721.7
OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 11
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 4.3
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.7
OTHER INFORMATION: EST HUMAN HIT: BE142985.1, EVALUATE 4.00e-81
OTHER INFORMATION: SWISSPROT HIT: Q13619, EVALUATE 8.00e-24
OTHER INFORMATION: NT HIT: AB012193.1, EVALUATE 4.00e-81
US-09-864-761-19055

Query Match 2.3%; Score 28.8; DB 9; Length 153;
Best Local Similarity 50.0%; Pred. No. 1e+03;
Matches 72; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 777 CATATCTTAGCAGGCCGACAGAACATCAGCTGAAGCATTTTACGAGATTGGTCAGAG 836
DB 7 CAGTCCCTGGGCTGTGCAAGACGCTGTGATTTAAAGTCCCAAGAGAAAGAGAGTG 66
QY 837 GAGGTGCTTCAACATGCGCTTTGTGCAAGATTAAAGATCTTTGTTGATTTTCAAG 896
DB 67 GAAGATGAGAGACAGATTCATTTTATGAGAGAGTTCAAGACAGATTTGTTAGATTAAG 126
QY 897 ATGAATAACATTGATGAGAAAGAA 920
DB 127 ATCAATCAATTCAGATGAAGAA 150

RESULT 15
US-10-242-535A-34626/c
Sequence 34626, Application US/10242535A
Publication No. US2004001363A1
GENERAL INFORMATION:
APPLICANT: ChondroGene Inc.
APPLICANT: Liew, C.C.
TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
FILE REFERENCE: 4231/2005
CURRENT APPLICATION NUMBER: US/10/242,535A
PRIOR FILING DATE: 2002-09-12
PRIOR APPLICATION NUMBER: US 10/085,783
PRIOR FILING DATE: 2002-02-28
PRIOR APPLICATION NUMBER: US 60/305,340
PRIOR FILING DATE: 2001-07-13
PRIOR APPLICATION NUMBER: US 60/275,017
PRIOR FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: US 60/271,955
PRIOR FILING DATE: 2001-02-28
NUMBER OF SEQ ID NOS: 58994
SOFTWARE: PatentIn version 3.2
SEQ ID NO 34626
LENGTH: 182
TYPE: DNA
ORGANISM: Human

US-10-242-535A-34626

Query Match 2.2%; Score 28.6; DB 17; Length 182;
Best Local Similarity 51.1%; Pred. No. 1.3e+03;
Matches 67; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 907 TTGATGAAAAGAACTGTGATCTTACAAAGAAAGCTGCTGATGATTTGAAAATTG 966
DB 173 TTTATGCTCGAAATGTGATCTTACAAATGAAAACAAATATGTTTAAATA 114
QY 967 AATCTTAGACTGCGTAGTAAAGTCTGAGAAAATTGAAAGCTGAGACCAAGTTA 1026
DB 113 AAGCAATAGGTTTGTAAACATTTGTTAATACCTTTGAAGGGTTATGATGAACATC 54
QY 1027 AATGCCCTTCT 1037
DB 53 TTTCTCTTTAT 43

Search completed: February 5, 2005, 11:55:46
Job time : 749 secs

This Page Blank (uspto)

QY	61	GATCCGCGGCGCCCCGGCTCCGACGGCGTCTTTCTTCAATATAAGAAACATGGGAAAC	120
Db	61	GATCCCGGCGCCCCCGCTCCGACGGCGTCTTTCTTCAATATAAGAAACATGGGAAAC	120
QY	121	TGATTACACACATTAGCTGATCATGGTGAAGATGTCACGTGCTGTCCTTCTCTTTCC	180
Db	121	TGATTACACACATTAGCTGATCATGGTGAAGATGTCACGTGCTGTCCTTCTCTTTCC	180
QY	181	TCTTGGCTACTTGTCTCTTGGACAAACAAATTCCGCTGACTCTGTACTGTACTTACTG	240
Db	181	TCTTGGCTACTTGTCTCTTGGACAAACAAATTCCGCTGACTCTGTACTGTACTTACTG	240
QY	241	AACTGCACATTCTGCATTGAAAGTTCAATCCCTATGCTGCACATGCGCTGTTCTGCC	300
Db	241	AACTGCACATTCTGCATTGAAAGTTCAATCCCTATGCTGCACATGCGCTGTTCTGCC	300
QY	301	CTTCAGACATATTTTGGCATCTGTTCAACAGATGATACACTGTCCCTATGAAATCTG	360
Db	301	CTTCAGACATATTTTGGCATCTGTTCAACAGATGATACACTGTCCCTATGAAATCTG	360
QY	361	AAATGACACAGATGCTGGCAGTGAACACGCTTAGTGGCAGCCCTGTGAGGGTTGCC	420
Db	361	AAATGACACAGATGCTGGCAGTGAACACGCTTAGTGGCAGCCCTGTGAGGGTTGCC	420
QY	421	AGTTTCCCCGACCTCCACGCTTGGCATCAGGGGCGACGTAGGAACTGTGTTTGT	480
Db	421	AGTTTCCCCGACCTCCACGCTTGGCATCAGGGGCGACGTAGGAACTGTGTTTGT	480
QY	481	GGATGACACAGTCATACAAATTATATAGATGTAGTGTAAAGATGGCTCTTGGCGG	540
Db	481	GGATGACACAGTCATACAAATTATATAGATGTAGTGTAAAGATGGCTCTTGGCGG	540
QY	541	CATGTGCATTTTCTCTTAATGGAAGCTTCTTTGTCACTGGCTCCTCATGTGGATTTAA	600
Db	541	CATGTGCATTTTCTCTTAATGGAAGCTTCTTTGTCACTGGCTCCTCATGTGGATTTAA	600
QY	601	CAGTGGGGATGATTAATATGAGGTCCTGCATAGAGAAAGAACATATCTTGGAAATTA	660
Db	601	CAGTGGGGATGATTAATATGAGGTCCTGCATAGAGAAAGAACATATCTTGGAAATTA	660
QY	661	CCTGCTGGGATTTTCTTCAACAGCCAGTTTCTGATGAGAAACAGGCTTCAGTTTTC	720
Db	661	CCTGCTGGGATTTTCTTCAACAGCCAGTTTCTGATGAGAAACAGGCTTCAGTTTTC	720
QY	721	GACTGGCATCATGTGTCAGATGTCGCAAGTCAAAATTGATGTTCTTTACCCATA	780
Db	721	GACTGGCATCATGTGTCAGATGTCGCAAGTCAAAATTGATGTTCTTTACCCATA	780
QY	781	TCTTAGCAAGGGCGCACAGACATCAAGCTGAAAGCAATTCACGAAATTTGTCAAGAGG	840
Db	781	TCTTAGCAAGGGCGCACAGACATCAAGCTGAAAGCAATTCACGAAATTTGTCAAGAGG	840
QY	841	TGCTCTCAACATGCGCTTTGTGCAAGATTTAAAGATCTTGTGTATTTTCAAGATGA	900
Db	841	TGCTCTCAACATGCGCTTTGTGCAAGATTTAAAGATCTTGTGTATTTTCAAGATGA	900
QY	901	ATAACATTGATGAAAGAACTGTTGAACTTTACAAAGAAAGTCTGGCTGATGATTTGA	960
Db	901	ATAACATTGATGAAAGAACTGTTGAACTTTACAAAGAAAGTCTGGCTGATGATTTGA	960
QY	961	AAATTGAATCTTACAGACTGCTGTATGAAGTCTGAGGAAATTTGMAAGCTCAGGACCA	1020
Db	961	AAATTGAATCTTACAGACTGCTGTATGAAGTCTGAGGAAATTTGMAAGCTCAGGACCA	1020
QY	1021	AGGTTAAATCCCTTCTTCAGGAAATTCGAAATGAATTAATATGTCATAACTGAGAAC	1080
Db	1021	AGGTTAAATCCCTTCTTCAGGAAATTCGAAATGAATTAATATGTCATAACTGAGAAC	1080
QY	1081	TTATGAAAGATCCGCTCATCGCATAGATGCTATTCATATAGAAAGAACCAATGAGAA	1140
Db	1081	TTATGAAAGATCCGCTCATCGCATAGATGCTATTCATATGAAAGAACCAATGAGAA	1140
QY	1141	ATTGATCAGCAAAAAGAACGTACAGTCCCATGACAAATCTGTGTTCTTCTTCACGCG	1200

Db	1141	ATTGATCAGCAAAAGAAAGTACAGTCCCATGACAAATCTTGTTCTTCTTCAGGG	1200
Qy	1201	TACTTACACCAATAGGACTCTGAAATATGCGCATATAGTGCCTGGAGACACCCAA	1260
Db	1201	TACTTACACCAATAGGACTCTGAAATATGCGCATATAGTGCCTGGAGACACCCAA	1260
Qy	1261	AGTAAAGAAATTC 1272	
Db	1261	AGTAAAGAAATTC 1272	
RESULT 2			
US-10-077-111-3			
; Sequence 3, Application US/10077111			
; Publication No. US20020187492A1			
; GENERAL INFORMATION:			
; APPLICANT: Toddard, C. Gordon			
; APPLICANT: Finger, Joshua N.			
; APPLICANT: Rillema, Jill			
; TITLE OF INVENTION: TEA			
; FILE REFERENCE: 3053-4114US2			
; CURRENT APPLICATION NUMBER: US/10/077,111			
; CURRENT FILING DATE: 2002-02-15			
; PRIOR APPLICATION NUMBER: 60/294,181			
; PRIOR FILING DATE: 2001-05-29			
; PRIOR APPLICATION NUMBER: 60/269,366			
; PRIOR FILING DATE: 2001-02-16			
; NUMBER OF SEQ ID NOS: 25			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 3			
; LENGTH: 1553			
; TYPE: DNA			
; ORGANISM: HUMAN			
US-10-077-111-3			
Query Match			
76.8%; Score 977.2; DB: 13; Length 1553;			
Best Local Similarity 81.9%; Pred. No. 1.9e-274;			
Matches 1265; Conservative 0; Mismatches 3; Indels 276; Gaps 1;			
Qy	1	GAATTCGGCTTTCACCTGCGCGGACGTCGACCCGACCCGCTGGGACCTTGAAGCG	60
Db	1	GAATTCGGCTTTCACCTGCGCGGACGTCGACCCGACCCGCTGGGACCTTGAAGCG	60
Qy	61	GATCCCGGCGCCCCCGCTCTGTGAGGCTGTTTCTTCAATGAAGAACGTGTAAC	120
Db	61	GATCCCGGCGCCCCCGCTCTGTGAGGCTGTTTCTTCAATGAAGAACGTGTAAC	120
Qy	121	TGATTCACACATTAGCTGATATATGTGACCATGTCACACTGCTGCGCTTCTCTTTCC	180
Db	121	TGATTCACACATTAGCTGATATATGTGACCATGTCACACTGCTGCGCTTCTCTTTCC	180
Qy	181	TCTTGGCTACTGCTGCTTGGACAAACAATTGCGCTGACTCGTTACGTACTTACTG	240
Db	181	TCTTGGCTACTGCTGCTTGGACAAACAATTGCGCTGACTCGTTACGTACTTACTG	240
Qy	241	AACGCCACATTCATTCATGTAAGTTTCATACCTATGCTGTCACCTGCTGTTCTGCC	300
Db	241	AACGCCACATTCATTCATGTAAGTTTCATACCTATGCTGTCACCTGCTGTTCTGCC	300
Qy	301	CTTCAGACATATTTTGGCAGCTGCTTCAACAGATGGTACCACTGCTCTTATGAAATCTG	360
Db	301	CTTCAGACATATTTTGGCAGCTGCTTCAACAGATGGTACCACTGCTCTTATGAAATCTG	360
Qy	361	AAATATGACATATGCTGGCAGTGTGATGTAAGACGCTTATGAGGACGCTGAGGGGTTGCC	420
Db	361	AAATATGACATATGCTGGCAGTGTGATGTAAGACGCTTATGAGGACGCTGAGGGGTTGCC	420
Qy	421	AGTTTTCGCCAGACTCCACGTTTGGCATAGGGGCGAGCTGATGGAATCTGTGTTTGT	480
Db	421	AGTTTTCGCCAGACTCCACGTTTGGCATAGGGGCGAGCTGATGGAATCTGTGTTTGT	480
Qy	481	GGAATGCACTATACAAATTAATATGATGTGTAAGTGTGTTAAAGATGCTCTTGGCGG	540

Db 481 GGATGACAGCATACAAATTTATATAGATGTGGAGTGAAGGCTCCCTGGCGG 540
Qy 541 CATGTGATTTTCCCTTAATGGAAGCTTTGTTCACATGCGTCTCTCATGTGTATTA 600
Db 541 CATGTGATTTTCCCTTAATGGAAGCTTTGTTCACATGCGTCTCTCATGTGTATTA 600
Qy 601 CAGTGTGGATGATTAATGAGTGTCTGATAGTGAAGAAAGCATGATCTTGAATTA 660
Db 601 CAGTGTGGATGATTAATGAGTGTCTGATAGTGAAGAAAGCATGATCTTGAATTA 660
Qy 661 CTTGCTGCGATTTTCTTCAAGCAGCTTTCTGATGAGAAAGGCTTTGATTTTC 720
Db 661 CTTGCTGCGATTTTCTTCAAGCAGCTTTCTGATGAGAAAGGCTTTGATTTTC 720
Qy 721 GACTGCGATCATGTGTGTCAGGATTTGCCAAGTCAAAATTTGATTTTCTTTACCA 780
Db 721 GACTGCGATCATGTGTGTCAGGATTTGCCAAGTCAAAATTTGATTTTCTTTACCA 780
Qy 781 TCTT----- 784
Db 781 TCTTAGTTTGAATTAATAAAGTACATGATGGGCACTGCTCTCTCTG 840
Qy 785 ----- 784
Db 841 CTTGTGCTTTTCCCGTATGGGAGATGCTAGTCTAGGCTCAGTGATAGTCTGCA 900
Qy 785 ----- 784
Db 901 TAGTATATGATTAATTAATGAGATTAATTAATTAATTAATTAATTAATTAAT 960
Qy 785 ----- 784
Db 961 TCACAACTGTGCTTTTGGACCTTAATCCTTTTACTGCTGCTTCAATGACAAA 1020
Qy 785 ----- 784
Db 1021 CAGTGAACATGCGCAATTTGACCTGGAACACTTTGCGAAGGCGCACAGAACATC 1080
Qy 805 AGCTGAAGCAATTTACCGAAGATTTGTCAGAGAGTGTCTCAACATGCTTTGTGAC 864
Db 1081 AGCTGAAGCAATTTACCGAAGATTTGTCAGAGAGTGTCTCAACATGCTTTGTGAC 864
Qy 865 AAGATTTAAAGATCTGTGTGTAATTTCAAGATGAATTAATGATGAGTGAAGAACTGT 924
Db 1141 AAGATTTAAAGATCTGTGTGTAATTTCAAGATGAATTAATGATGAGTGAAGAACTGT 924
Qy 925 TGAATCTTACAAAGAAAGTGTGCTGATGATTTGAAATGAAATCTCTAGACTGCGTA 984
Db 1201 TGAATCTTACAAAGAAAGTGTGCTGATGATTTGAAATGAAATCTCTAGACTGCGTA 984
Qy 985 GTAAAGTCTGAGGAAATGGAAGTCAAGACCAAGGTTAAATCCCTTTCTTCAAGAA 1044
Db 1261 GTAAAGTCTGAGGAAATGGAAGTCAAGACCAAGGTTAAATCCCTTTCTTCAAGAA 1044
Qy 1045 TTCTGATGATTTATATGTCATTAATTAAGAACTTAATGAAGATCCGCTCATGCGAT 1104
Db 1321 TTCTGATGATTTATATGTCATTAATTAAGAACTTAATGAAGATCCGCTCATGCGAT 1104
Qy 1105 CAGATGCTATTTCAATGAAAGAAAGCAATGAAATTTGATGAGAAAGAAAGCTTA 1164
Db 1381 CAGATGCTATTTCAATGAAAGAAAGCAATGAAATTTGATGAGAAAGAAAGCTTA 1164
Qy 1165 CAAATGCCATGAAATCTTGTCTTCTTCAAGGCTAATTAACCAATGAGACTGCA 1224
Db 1441 CAAATGCCATGAAATCTTGTCTTCTTCAAGGCTAATTAACCAATGAGACTGCA 1224
Qy 1225 AAATGCGCATTAATGATGCTGAGACACACCAAAAGTAAAGA 1288
Db 1501 AAATGCGCATTAATGATGCTGAGACACACCAAAAGTAAAGA 1288
Qy 1501 AAATGCGCATTAATGATGCTGAGACACACCAAAAGTAAAGA 1544

US-10-077-111-1
; Sequence 1, Application US/10077111
; Publication No. US20020187492A1
; GENERAL INFORMATION:
; APPLICANT: Toddard, C. Gordon
; APPLICANT: Finger, Joshua N.
; APPLICANT: Rillema, Jill
; TITLE OF INVENTION: TBA
; FILE REFERENCE: 1053-4114US2
; CURRENT APPLICATION NUMBER: US/10/077, 111
; CURRENT FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 60/294,181
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: 60/269,366
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1818
; TYPE: DNA
; ORGANISM: HUMAN
US-10-077-111-1
Query Match 75.9%; Score 965.8; DB 13; Length 1818;
Best Local Similarity 81.8%; Pred. No. 4.5e-271;
Matches 1253; Conservative 0; Mismatches 2; Indels 276; Gaps 1;
Qy 12 TCACCTGCGGCGACGTCAGCCGACCCGCTGGGCACTTTGAAGCGGATCCCGCG 71
Db 49 TCACCTGCGGCGACGTCAGCCGACCCGCTGGGCACTTTGAAGCGGATCCCGCG 108
Qy 72 CCCCCGCTCTGACGAGGCTTTTCTTAATTAAGAAAGTGAAGAACTGATTCACACA 131
Db 109 CCCCCGCTCTGACGAGGCTTTTCTTAATTAAGAAAGTGAAGAACTGATTCACACA 168
Qy 132 TTAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 191
Db 169 TTAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 228
Qy 192 TGCTCTTGAACAAACAAATTCGCTGTACTGCTGTACTGCTGTACTGCTGTACT 251
Db 229 TGCTCTTGAACAAACAAATTCGCTGTACTGCTGTACTGCTGTACTGCTGTACT 288
Qy 252 TTTCCATTAAGATTTTCACTTAATGCTGATGCTGATGCTGATGCTGATGCTGAT 311
Db 289 TTTCCATTAAGATTTTCACTTAATGCTGATGCTGATGCTGATGCTGATGCTGAT 348
Qy 312 ATTTGGCATGCTGCTCAAGATGATGATGATGATGATGATGATGATGATGATGAT 371
Db 349 ATTTGGCATGCTGCTCAAGATGATGATGATGATGATGATGATGATGATGATGAT 408
Qy 372 ATGCTGACATGATGAAACAGCTGATGAGACCCCTGTGAGGTTTCCAGTTTCCCA 431
Db 409 ATGCTGACATGATGAAACAGCTGATGAGACCCCTGTGAGGTTTCCAGTTTCCCA 468
Qy 432 GACTCCAGTGTGTGATGATGAGGCGATGATGAGAACTGATGATGATGATGATGAT 491
Db 469 GACTCCAGTGTGTGATGATGAGGCGATGATGAGAACTGATGATGATGATGATGAT 528
Qy 492 TCATCAAAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 551
Db 529 TCATCAAAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 588
Qy 552 TTTCTTAATGAAAGCTTTTGTGATGATGATGATGATGATGATGATGATGATGAT 611
Db 589 TTTCTTAATGAAAGCTTTTGTGATGATGATGATGATGATGATGATGATGATGAT 648
Qy 612 GATTAATGAGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 671
Db 649 GATTAATGAGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 708
Qy 672 TTTTCTCAAGCAGGCTTTCTGATGAGAAAGGCTTCAAGTTTTCAGCTGACATCA 731

Db 709 TTTCTTCACAGCCAGTTTCGTGAGAGAAAGAGCTTCAGTTTTTTCAGTCGGACATCA 768
Qy 732 TGTGTCAGAGATTGGCCAAAGTCAAAATTTGGATTGTTTCTTTTACCCATATCTT----- 784
Db 769 TGTGTCAGAGATTGGCCAAAGTCAAAATTTGGATTGTTTCTTTTACCCATATCTTAGGTTT 828
Qy 785 ----- 784
Db 829 GAATTAATAATATAAGTACACTGAGTGGGACACTGTGCTCCTGTTTGTGCTGTGCTTTT 888
Qy 785 ----- 784
Db 889 TCCCATGATGAGGAGATGATCTAGTCTCAGGGTCAGTGTAAAGTCTGTCTGTATATATGAT 948
Qy 785 ----- 784
Db 949 ACTAATATCTAGAATATATCTTACACATTTGACTCAGACACACAGGTATGTACAACTTGT 1008
Qy 785 ----- 784
Db 1009 GCTTTTGACACTAATACCTTTTACTGTTACTGTCTCAATGTGACAAAGACATGACATC 1068
Qy 785 -----AGCAAGGCGCACAAACATCAGCTGAAACAA 815
Db 1069 TGGCAATTTGACCTGGAACACTTTGCGCAAGCAAGGCGCACAGAAACATCAGCTGAAACAA 1128
Qy 816 TTACCGAAGATTGGTCAGAGAGAGTGTCTCAACATGAGCTTTGTGACAAAGATTTTAA 875
Db 1129 TTTTACCGAAGATTGGTCAGAGAGAGTGTCTCAACATGAGCTTTGTGACAAAGATTTTAA 1188
Qy 876 GATCTGTGTTGATTTTCAAGATGAATTAACATTTGATGAAAGAAAGAACTGTTGAATCTTACA 935
Db 1189 GATCTGTGTTGATTTTCAAGATGAATTAACATTTGATGAAAGAAAGAACTGTTGAATCTTACA 1248
Qy 936 AAAGAAAGTGTGCTGATGATTTGAAATTTGAATCTCTAGAGATGCGTATGAAAGTCTG 995
Db 1249 AAAGAAAGTGTGCTGATGATTTGAAATTTGAATCTCTAGAGATGCGTATGAAAGTCTG 1308
Qy 996 AGGAAATTTGAAGAGCTCAGAGACCAAGTTAAATCCCTTTCTTCAGGAAATTCCTGATGAA 1055
Db 1309 AGGAAATTTGAAGAGCTCAGAGACCAAGTTAAATCCCTTTCTTCAGGAAATTCCTGATGAA 1368
Qy 1056 TTTATATGTCCAATTAACATGAGAACTTATGAAAGATCCGTCATCGCATGAGATGCGTAT 1115
Db 1369 TTTATATGTCCAATTAACATGAGAACTTATGAAAGATCCGTCATCGCATGAGATGCGTAT 1428
Qy 1116 TCATATGAAAAAGAGCAATGAAAAATTTGATCAGCAAAAAAGAAAGCTACAGTCCCATG 1175
Db 1429 TCATATGAAAAAGAGCAATGAAAAATTTGATCAGCAAAAAAGAAAGCTACAGTCCCATG 1488
Qy 1176 ACAAAATCTGTGTTCTCTTCAGGCGGTACTTACACCAATAGSACTCTGAAATGGCCATC 1235
Db 1489 ACAAAATCTGTGTTCTCTTCAGGCGGTACTTACACCAATAGSACTCTGAAATGGCCATC 1548
Qy 1236 AATAGATGCTGAGACACACCAAAAGTAAA 1266
Db 1549 AATAGATGCTGAGACACACCAAAAGTAAA 1579

RESULT 4
US-09-971-392-207
; Sequence 207, Application US/09971392
; Publication No. US20030134283A1
; GENERAL INFORMATION:
; APPLICANT: Peterson, David P.
; APPLICANT: Peterson, Cecilia I.
; APPLICANT: Cooke, Benjamin G.
; TITLE OF INVENTION: GENES REGULATED IN DENDRITIC CELL DIFFERENTIATION
; FILE REFERENCE: PA-0029 US
; CURRENT APPLICATION NUMBER: US/09/971,392
; PRIORITY FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: 60/237,652
; PRIOR FILING DATE: 2000-10-03

; NUMBER OF SEQ ID NOS: 260
; SOFTWARE: PERL Program
; SEQ ID NO 207
; LENGTH: 1817
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Template ID: 158923.9
US-09-971-392-207

Query Match 75.8%; Score 964.2; DB 10; Length 1817;
Best Local Similarity 81.8%; Pred. No. 1.3e-270;
Matches 1252; Conservative 0; Mismatches 3; Indels 276; Gaps 1;

Qy 12 TCACCTGCGGAGACGTGACCCGACCCGCGTGGGACCTTGAAGGGGAGATCCCGCGG 71
Db 49 TCACCTGCGGAGACGTGACCCGACCCGCGTGGGACCTTGAAGGGGAGATCCCGCGG 108
Qy 72 CCCCCTCTCTGACGAGCTGTTTTTCTTCAATTAAGAACATGATGAACTGATTCACACA 131
Db 109 CCCCCTCTCTGACGAGCTGTTTTTCTTCAATTAAGAACATGATGAACTGATTCACACA 168
Qy 132 TTAGCTGATCATGTGACGATGTCAACTGCTGCTTCTCTTTTCCCTTTGCTACT 191
Db 169 TTAGCTGATCATGTGACGATGTCAACTGCTGCTTCTCTTTTCCCTTTGCTACT 228
Qy 192 TGCTCCTTGGACAAACAAATTTGGCCCTGTACTGTTACGATCTTACTGAACTGCCACAT 251
Db 229 TGCTCCTTGGACAAACAAATTTGGCCCTGTACTGTTACGATCTTACTGAACTGCCACAT 288
Qy 252 TCTCCATTGAAGTTTCTATCTATGCTGTCTCACTGCTGCTTCTCCCTTCAGGACAT 311
Db 289 TCTCCATTGAAGTTTCTATCTATGCTGTCTCACTGCTGCTTCTCCCTTCAGGACAT 348
Qy 312 ATTTTGGCATCGTGTTCACAGATGATGACACTGTCTTATGGAATCTGAAATGAGACAG 371
Db 349 ATTTTGGCATCGTGTTCACAGATGATGACACTGTCTTATGGAATCTGAAATGAGACAG 408
Qy 372 ATGCTGGCAGTATGGAACAGCTATGAGCAGCCCTGTAGAGGTTGCCAGTTTCCCGCA 431
Db 409 ATGCTGGCAGTATGGAACAGCTATGAGCAGCCCTGTAGAGGTTGCCAGTTTCCCGCA 468
Qy 432 GACTCAGCTGTTTGGCATCAGGGGACGCTGATGGAATGATGTTTGTGAATGACAG 491
Db 469 GACTCAGCTGTTTGGCATCAGGGGACGCTGATGGAATGATGTTTGTGAATGACAG 528
Qy 492 TCATACAAATTAATATGATGTGTGATGTTAAAGATGCTCCTTGGCGCATGTGCATTT 551
Db 529 TCATACAAATTAATATGATGTGTGATGTTAAAGATGCTCCTTGGCGCATGTGCATTT 588
Qy 552 TCTCCTAATGGAAGCTTCTTTTGTCACTGGCTCTCACTGAGTATTAACAGTGGGAT 611
Db 589 TCTCCTAATGGAAGCTTCTTTTGTCACTGGCTCTCACTGAGTATTAACAGTGGGAT 648
Qy 612 GATAAATGAGGTGTCTGATGATGAAAAAGCAGATGTTGGAATTAACCTGCTCGAT 671
Db 649 GATAAATGAGGTGTCTGATGATGAAAAAGCAGATGTTGGAATTAACCTGCTCGAT 708
Qy 672 TTTTCTTCACAGCAGCTTCTGTATGAGAAACAAAGTCTTCAAGTTTGTGACTGCGATCA 731
Db 709 TTTTCTTCACAGCAGCTTCTGTATGAGAAACAAAGTCTTCAAGTTTGTGACTGCGATCA 768
Qy 732 TGTGTCAGAGATTGGCCAAAGTCAAAATTTGGATTGTTTCTTTTACCCATATCTTAGGTTT 784
Db 769 TGTGTCAGAGATTGGCCAAAGTCAAAATTTGGATTGTTTCTTTTACCCATATCTTAGGTTT 828
Qy 785 ----- 784
Db 829 GAATTAATAATATAAGTACACTGAGTGGGACACTGTGCTCCTGTTGTGCTGTGCTTTT 888
Qy 785 ----- 784

Db 889 TCCCATGATGGGAGATGCTAGTCTCAGAGGTCAAGTGATTAAGTCTGCTATGATATGAT 948
Qy 785 ----- 784
Db 949 ACTAATACGTGAATATATCTTACACATTTAGCTACGACACACAGGTATGTCAAACTTGT 1008
Qy 785 ----- 784
Db 1009 GCTTTGCACTTAATACCTTTTACTTGTCTAGTGTTCATATGACAAAACAGTGAACATC 1068
Qy 785 -----ACCAAGCGCACAGAACATCAGCTGAAGCA 815
Db 1069 TGGCAATTTGACCTGGAACACATTTGCGCAAGAGACAGAACATCAGCTGAAGCA 1128
Qy 816 TTTTACCGAAGTTGGTCAAGAGAGTGTCTCAACATGAGCTTTTGTGCAAGATTTTAA 875
Db 1129 TTTTACCGAAGTTGGTCAAGAGAGTGTCTCAACATGAGCTTTTGTGCAAGATTTTAA 1188
Qy 876 GATCTGTGTGTTATTTTCAAGATGATTAATCAATTTGATGGAAGAACTGTGATCTTACA 935
Db 1189 GATCTGTGTGTTATTTTCAAGATGATTAATCAATTTGATGGAAGAACTGTGATCTTACA 1248
Qy 936 AAAGAAAGTGTGCTGATGATTTGAAAATTGAATCTTGAAGCTGCTGATGAAGTGTG 995
Db 1249 AAAGAAAGTGTGCTGATGATTTGAAAATTGAATCTTGAAGCTGCTGATGAAGTGTG 1308
Qy 996 AGGAAATTTGAAGCTCAGACCAAGTTAAATCCCTTTTCAAGAAATTCCTGATGAA 1055
Db 1309 AGGAAATTTGAAGCTCAGACCAAGTTAAATCCCTTTTCAAGAAATTCCTGATGAA 1368
Qy 1056 TTTTATGTCCTCAATTAAGTGAAGCTTGAAGAGTCCGATCATGCTGATGAGTGTAT 1115
Db 1369 TTTTATGTCCTCAATTAAGTGAAGCTTGAAGAGTCCGATCATGCTGATGAGTGTAT 1428
Qy 1116 TCATATGAAAAGGAAGCAATGAAAATTGATCAGCAAAAAGAAACGTACAGTCCCATG 1175
Db 1429 TCATATGAAAAGGAAGCAATGAAAATTGATCAGCAAAAAGAAACGTACAGTCCCATG 1488
Qy 1176 ACAATCTGTGTTCTTCTTCAAGCGGTACTTACACCAATTAGAAGCTGAAAATGCGCATC 1235
Db 1489 ACAATCTGTGTTCTTCTTCAAGCGGTACTTACACCAATTAGAAGCTGAAAATGCGCATC 1548
Qy 1236 AATGATGCGTGAAGACACACCAAAAGTAAA 1266
Db 1549 AATGATGCGTGAAGACACACCAAAAGTAAA 1579

RESULT 5

US-10-037-270-768
; Sequence 768, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyen
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yundong
; APPLICANT: Wang, Duncui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Tillinghaast, John
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/10/037,270

; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pc_Fl_genes Version 1.0
; SEQ ID NO 768
; LENGTH: 1844
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (164)..(1594)
US-10-037-270-768
Query Match 75.8%; Score 964.2; DB 15; Length 1844;
Best Local Similarity 81.8%; Pred. No. 1.3e-270;
Matches 1252; Conservative 0; Mismatches 3; Indels 276; Gaps 1;
Qy 12 TCACCTGCGGCGACGTCGACCCGACCGCCGCGGACCTTGAAAGCGGATCCCGGCG 71
Db 65 TCACCTGCGGCGACGTCGACCCGACCGCCGCGGACCTTGAAAGCGGATCCCGGCG 124
Qy 72 CCCCCGCTCGACGCGTGTCTTCAATTAAGAAACATGGAATGATTCACACA 131
Db 125 CCCCCGCTCGACGCGTGTCTTCAATTAAGAAACATGGAATGATTCACACA 184
Qy 132 TTAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 191
Db 185 TTAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 244
Qy 192 TGCTCTTGAACAAACATTCGCTGATCTGATGATGATGATGATGATGATGATGATGAT 251
Db 245 TGCTCTTGAACAAACATTCGCTGATCTGATGATGATGATGATGATGATGATGATGAT 304
Qy 252 TCTCATTGAAGTTTCACTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 311
Db 305 TCTCATTGAAGTTTCACTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 364
Qy 312 ATTTGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 371
Db 365 ATTTGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 424
Qy 372 ATGCTGCGATGATGAAACGCTTATGAGGAGCCCTGTGAGGCTTCCGATTTCCCA 431
Db 425 ATGCTGCGATGATGAAACGCTTATGAGGAGCCCTGTGAGGCTTCCGATTTCCCA 484
Qy 432 GACTCCAGTGTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 491
Db 485 GACTCCAGTGTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 544
Qy 492 TCATCAAAATTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 551
Db 545 TCATCAAAATTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 604
Qy 552 TCTCTAATGAAGCTTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 611
Db 605 TCTCTAATGAAGCTTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 664
Qy 612 GATTAATGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 671
Db 665 GATTAATGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 724
Qy 672 TTTTCTTCAAGCAGTTTCTGATGAGAAACAAGTCTTCAAGTCTTCTTCTTCTTCTTCT 731
Db 725 TTTTCTTCAAGCAGTTTCTGATGAGAAACAAGTCTTCAAGTCTTCTTCTTCTTCTTCT 784
Qy 732 TGTGTCAGGATTCGACAGTCAAAATTTGATGATGATGATGATGATGATGATGATGAT 784
Db 785 TGTGTCAGGATTCGACAGTCAAAATTTGATGATGATGATGATGATGATGATGATGAT 844
Qy 785 ----- 784

Db 845 GAATTAAATATAAGTACAGTGAAGGAGCTGCTCTCTGCTTGTGCTTGTGCTTTT 904
Qy 785 ----- 784
Db 905 TCCCATATGAGGACAGATGCTAGTCTCAGGGCTCAGTGATTAAGTGTCTATATATATGAT 964
Qy 785 ----- 784
Db 965 ACTAATACAGAAATATACTTACACATGACACAGAGTATGTACAACTTGT 1024
Qy 785 ----- 784
Db 1025 GCTTTTGACCTAATACCTTTTACTGTCTAAGTTCATGAGCAAAACAGTGACATC 1084
Qy 785 -----AGCAAGGCGACAGAAACATCAGCTGAAGCA 815
Db 1085 TGGCAATTGACCTGAGAAACCTTTGCCAGAGAGACAGAACATCAGCTGAAGCA 1144
Qy 816 TTACCGAAGATTGCTCAGAGAGAGTCTCTCAACATGCTTTGTGACAAAGATTTTAAA 875
Db 1145 TTATCCGAAGTTGCTCAGAGAGAGATGTCTCAACATGCTTTGTGACAAAGATTTTAAA 1204
Qy 876 GATCTTTGTGATTTTCAAGATGATTAACATTGATGAGAAAGAACTGTTGAATCTTACA 935
Db 1205 GATCTTTGTGATTTTCAAGATGATTAACATTGATGAGAAAGAACTGTTGAATCTTACA 1264
Qy 936 AAGAAAGTCTGCTGATGATTTGAAATTTGAATCTCTAAGGACGCGTGAAGAAAGTCTG 995
Db 1265 AAGAAAGTCTGCTGATGATTTGAAATTTGAATCTCTAAGGACGCGTGAAGAAAGTCTG 1324
Qy 996 AGGAAATTTGAGAGCTCAGAGACCAAGTTAAATCCCTTCTTCAAGAAATTCCTGATGA 1055
Db 1325 AGGAAATTTGAGAGCTCAGAGACCAAGTTAAATCCCTTCTTCAAGAAATTCCTGATGA 1384
Qy 1056 TTTTATATGCTCAATTAACAGAACTTATGAAAGATCCGGTCATGCGATCAGATGCGTAT 1115
Db 1385 TTTTATATGCTCAATTAACAGAACTTATGAAAGATCCGGTCATGCGATCAGATGCGTAT 1444
Qy 1116 TCATATGAAAGAGAAAGCAATGAGAAATTTGATCAGAAAAAGAAAGCAAGTCCCATG 1175
Db 1445 TCATATGAAAGAGAAAGCAATGAGAAATTTGATCAGAAAAAGAAAGCAAGTCCCATG 1504
Qy 1176 ACAAAATCTTGTCTTCTTCCCTTACGCGTACTTACACCAATATGAGACTTGAAATGCGCATC 1235
Db 1505 ACAAAATCTTGTCTTCTTCCCTTACGCGTACTTACACCAATATGAGACTTGAAATGCGCATC 1564
Qy 1236 AATATGATGCTGAGAGACACCAAAAGTAAA 1266
Db 1565 AATATGATGCTGAGAGACACCAAAAGTAAA 1595

RESULT 6
US-10-117-722-768
; Sequence 768, Application US/10117722
; Publication No. US20030219744A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Adundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030219744A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2BCIP
; CURRENT APPLICATION NUMBER: US/10/117, 722
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21

; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pc_fl_genes Version 1.0
; SEQ ID NO 768
; LENGTH: 1844
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (164) ..(1594)
US-10-117-722-768

Query Match 75.8%; Score 964.2; DB 17; Length 1844;
Best Local Similarity 81.8%; Prid. No. 1.3e-270;
Matches 1252; Conservative 0; Mismatches 3; Indels 276; Gaps 1;

Qy 12 TCACCTGCGGAGACGAGACCGGACCGCGGACCTTGAAGCGGATCCGCGG 71
Db 65 TCACCTGCGGAGACGAGACCGGACCGCGGACCTTGAAGCGGATCCGCGG 124
Qy 72 CCCCCGCTCTGAGGCTGTTTTCTTCAATTAAGAAACATGAGTGAATTCACACA 131
Db 125 CCCCCGCTCTGAGGCTGTTTTCTTCAATTAAGAAACATGAGTGAATTCACACA 184
Qy 132 TTAGCTGATATGTAAGATGTAACATGCTGTGCTCTCTCTTTCCCTTGGCTACT 191
Db 185 TTAGCTGATATGTAAGATGTAACATGCTGTGCTCTCTCTTTCCCTTGGCTACT 244
Qy 192 TGCTCCCTGGAACAAACATTTGCGCTGTACTGTGTAAGTACTTGAACCTGACAT 251
Db 245 TGCTCCCTGGAACAAACATTTGCGCTGTACTGTGTAAGTACTTGAACCTGACAT 304
Qy 252 TCTCATTTGAAGTTTCAATCTATGCTGTCACTGCTGTGTTCTCCCTTCAGACAT 311
Db 305 TCTCATTTGAAGTTTCAATCTATGCTGTCACTGCTGTGTTCTCCCTTCAGACAT 364
Qy 312 ATTTTGGCATGTTTCAACAGATGTAACATGCTGTCTTATGAAATCTGAAATGACAG 371
Db 365 ATTTTGGCATGTTTCAACAGATGTAACATGCTGTCTTATGAAATCTGAAATGACAG 424
Qy 372 ATGCTGGAGTATGTAAGAGCTAGTGGAGGCTGTGAGGTTGGCAGTTTCCCA 431
Db 425 ATGCTGGAGTATGTAAGAGCTAGTGGAGGCTGTGAGGTTGGCAGTTTCCCA 484
Qy 432 GACTCCACGTTTGGCATCAGGGGACGTGATGAACTGTGTTTGTGAAATGACAG 491
Db 485 GACTCCACGTTTGGCATCAGGGGACGTGATGAACTGTGTTTGTGAAATGACAG 544
Qy 492 TCATCAAAATTAATATGATGTGTATGTTAAAGATGGCTCCTTGGCGGATGTGATT 551
Db 545 TCATCAAAATTAATATGATGTGTATGTTAAAGATGGCTCCTTGGCGGATGTGATT 604
Qy 552 TCTCCTAATGGAAGCTTCTTGTCACTGGCTCCTCAATGTGTGATTTAACAGTGGAGT 611
Db 605 TCTCCTAATGGAAGCTTCTTGTCACTGGCTCCTCAATGTGTGATTTAACAGTGGAGT 664
Qy 612 GATTAATGAGTGTGTGATGTAAGAAAGCAATGATCTTGAATTAACCTGCTGAT 671
Db 665 GATTAATGAGTGTGTGATGTAAGAAAGCAATGATCTTGAATTAACCTGCTGAT 724
Qy 732 TGTGTGAGATTTGCCAAGTCAAAATTTGATGTTTCTTTTCCCATATCTT----- 784
Db 785 TGTGTGAGATTTGCCAAGTCAAAATTTGATGTTTCTTTTCCCATATCTTAAAGTTT 844
Qy 785 ----- 784
Db 845 GAATTAAATATAAGTACAGTGAAGGAGCTGCTCTCTGCTTGTGCTTGTGCTTTT 904
Qy 785 ----- 784

Db 905 TCCCATGATGGGAGATGCTAGTCTCAGGGTCAGTGATAGTCTGTATATATGAT 964
Qy 785 ----- 784
Db 965 ACTAATATCGAATATATATCTTACACATTTGACTCAGACACACAGGATATGTCACACTTGT 1024
Qy 785 ----- 784
Db 1025 GCTTTGACCTAATACCCCTTTTACTTGTCTAGTGTTCATATGACAAAACAGTGAATC 1084
Qy 785 -----ACGAGGCGCACAGAACTCAGCTGAGCAA 815
Db 1085 TGGCAATTGACCTGGAACAATTGCGCAAGCAGAGCAGACAGAACTCAGCTGAAACAA 1144
Qy 816 TTTACCGAAGTGGTGTGAGAGGAGTGTCTCAATGCGCTTGTGACAAAGATTAA 875
Db 1145 TTTACCGAAGTGGTGTGAGAGGAGTGTCTCAATGCGCTTGTGACAAAGATTAA 1204
Qy 876 GATCTGTGTGATTTTCAAGATGAATTAATCATTTGATGGAAGAAAGAACTGTTGAATCTTACA 935
Db 1205 GATCTGTGTGATTTTCAAGATGAATTAATCATTTGATGGAAGAAAGAACTGTTGAATCTTACA 1264
Qy 936 AAAGAAAGTGTGCTGATGATTTGAAAATGGAATCTCTAGACTGCGTAAAGTGTG 995
Db 1265 AAAGAAAGTGTGCTGATGATTTGAAAATGGAATCTCTAGACTGCGTAAAGTGTG 1324
Qy 996 AGGAAATTTGAAGGCTCAGGACCAAGTTAAATCCCTTTTCTAGGAATTCCTGATGAA 1055
Db 1325 AGGAAATTTGAAGGCTCAGGACCAAGTTAAATCCCTTTTCTAGGAATTCCTGATGAA 1384
Qy 1056 TTTATATGTCCTCAATTAATCAGAACTTATGAAAGATCCGGTCATGCGATCAGATGCGTAT 1115
Db 1385 TTTATATGTCCTCAATTAATCAGAACTTATGAAAGATCCGGTCATGCGATCAGATGCGTAT 1444
Qy 1116 TCATATGAAAAAGAGCATGGAATTTGATTCAGCAAAAAGAAACGTACAACTGCCATG 1175
Db 1445 TCATATGAAAAAGAGCATGGAATTTGATTCAGCAAAAAGAAACGTACAACTGCCATG 1504
Qy 1176 ACAATATCTTGTCTTCTTCTTCAAGGCTTACACCAATATGAGATCTTGAATATGCCATC 1235
Db 1505 ACAATATCTTGTCTTCTTCTTCAAGGCTTACACCAATATGAGATCTTGAATATGCCATC 1564
Qy 1236 AATGATGGCTGGAGACACACCAAAAGTAA 1266
Db 1565 AATGATGGCTGGAGACACACCAAAAGTAA 1595

RESULT 7
US-10-287-218-40
; Sequence 40, Application US/10287218
; Publication No. US20030198975A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: AZIMZAI, Yalda; AU-YOUNG, Janice K.
; APPLICANT: BARRA, Sajeev; BAUGHN, Mariah R.
; APPLICANT: BECHA, Shanya D.; BOROMSKY, Mark L.
; APPLICANT: BUFORD, Neil; DING, Li
; APPLICANT: ELDIOTT, Vicki S.; EMERLING, Brooke M.
; APPLICANT: GANDHI, Ameena R.; GIERZEN, Kimberly J.
; APPLICANT: GRIFFIN, Jennifer A.; HAPALIA, April J.A.
; APPLICANT: HONCHELL, Cynthia D.; LAL, Preeti G.
; APPLICANT: LEE, Soo Yeun; LU, Dying Aina M.
; APPLICANT: ARVIZU, Chandra S.; RAMKUMAR, Jayalaxmi
; APPLICANT: REDDY, Roopa; SANJANMALA, Madhu, M.
; APPLICANT: TANG, Y. Tom; WALIA, Nardier K.
; APPLICANT: WANG, Yu-mei; E; WARREN, Bridget A.
; APPLICANT: XU, Yuming; YANG, Junning
; APPLICANT: YAO, Monique G.; YDE, Henry
; APPLICANT: ZEBARADIAN, Yeganeh
; TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH
; FILE REFERENCE: PI-0417 USA
; CURRENT APPLICATION NUMBER: US/10/287,218
; CURRENT FILING DATE: 2002-10-31

; PRIOR APPLICATION NUMBER: PCT/US02/11152
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: US 60/349,705
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: US 60/295,263
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/295,340
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/293,727
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/291,946
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/291,662
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: US 60/287,228
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/286,820
; PRIOR FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: US 60/283,294
; PRIOR FILING DATE: 2001-04-11
; Remaining Prior Application data removed - See File Wrapper or PAM.
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 40
; LENGTH: 1773
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030198975A1 7483131CB1
US-10-287-218-40

Query Match 70.3%; Score 894; DB 16; Length 1773;
Best Local Similarly 80.6%; Pred. No. 4,4e-250;
Matches 1186; Conservative 0; Mismatches 10; Indels 276; Gaps 1;

Qy 71 GCCCGCGCTCCGACAGGCTTTTCTTCAATTAAGAAATGATGTAATTCATGAC 130
Db 20 GCGCGGTGGCGGAGGCTTTTCTTCAATTAAGAAATGATGTAATTCATGAC 79
Qy 131 ATTAGTGATCATGTGACGATGTCATGCTGCTGCTTCTTCTTCTTCTGCTGAC 190
Db 80 ATTAGTGATCATGTGACGATGTCATGCTGCTGCTTCTTCTTCTTCTGCTGAC 139
Qy 191 TTGCTCTTGGACAAAACAATTGCTGTACTGTGTTACGTGACTTACTGAACTGCCACA 250
Db 140 TTGCTCTTGGACAAAACAATTGCTGTACTGTGTTACGTGACTTACTGAACTGCCACA 199
Qy 251 TTCTCCATTAAGTTTCATCTATGCTGTCACCTGCTGCTGTTTCTCCCTTCAGACA 310
Db 200 TTCTCCATTAAGTTTCATCTATGCTGTCACCTGCTGCTGTTTCTCCCTTCAGACA 259
Qy 311 TATTGGCATCGTGTTCACACAGATGGTACCACTGTCCTTATGAAATCTGAAATGAGACA 370
Db 260 TATTGGCATCGTGTTCACACAGATGGTACCACTGTCCTTATGAAATCTGAAATGAGACA 319
Qy 371 GATCTGCGAGTATGAGAACAGCTTATGTCGACGCTTGTAGAGGCTTTCAGATTTTCCC 430
Db 320 GATCTGCGAGTATGAGAACAGCTTATGTCGACGCTTGTAGAGGCTTTCAGATTTTCCC 379
Qy 431 AGACTCCAGCTGTTTGGCATCAGGGGCAAGTATGAACTGTGTTTGTGAAATGAGACA 490
Db 380 AGACTCCAGCTGTTTGGCATCAGGGGCAAGTATGAACTGTGTTTGTGAAATGAGACA 439
Qy 491 GTCATACAAATTAATGATGTGTAGTGTAAAGATGCTCTTGGCGGCAATGTGAT 550
Db 440 GTCATACAAATTAATGATGTGTAGTGTAAAGATGCTCTTGGCGGCAATGTGAT 499
Qy 551 TTCTCTTAATGAGAGCTTCTTGTCACTGCTCTTCAATGTGTGATTTAAACAGTGGGA 610
Db 500 TTCTCTTAATGAGAGCTTCTTGTCACTGCTCTTCAATGTGTGATTTAAACAGTGGGA 559
Qy 611 TGATTAATGAGAGTGTGTGATGTGTAAGAAAGACATGATCTTGAATTAACCTGCTGGA 670

```

Db      560 TGAATAAATAGAGTGTCTGATAGTGAATAAGCAACATGATCTGGAAATTCCTGCTCGGA 619
QY      671 TTTTCTTCACAGCAGATTCTGATGGAGAACAAAGCTTCAGTTTTCGACGGCAGTC 730
Db      620 TTTTCTTCACAGCAGATTCTGATGGAGAACAAAGCTTCAGTTTTCGACGGCAGTC 679
QY      731 ATGTGTGACAGATTGCCAATGTCAAAATTTGGATTGTTCTTTTACCATACTCTT----- 784
Db      680 ATGTGTGACAGATTGCCAATGTCAAAATTTGGATTGTTCTTTTACCATACTCTTAGGTTT 739
QY      785 ----- 784
Db      740 TGAATTAATAATATAAAGTACACTGAGTGGGCACTGTGCTCTGTCTGCTGTGCTT 799
QY      785 ----- 784
Db      800 TTCCCATGATGGCAGATGCTAGTCTCAGGGTCAGTGGAATAGTGTGATATATATGA 859
QY      785 ----- 784
Db      860 TACTAATACTGAGAATATCTTCAACATTTGACTCAGACACCAAGTATGTCAACTG 919
QY      785 ----- 784
Db      920 TGTCTTTCACCTAATACCTCTTTTACTGCTACTGTTCAATGGACAAACAGTGAACAT 979
QY      785 ----- 784
Db      980 CTGGCAATTTGACCTGGAAAACCTTTGCCAAGGAGACAGAACATATAGCTGAAGCA 1039
QY      815 ATTTACGGAAGATTGTGTGACAGAGAGTGTCTCAACATGCTTGTGTCACAAAGATTAA 874
Db      1040 ATTTACGGAAGATTGTGTGACAGAGAGTGTCTCAACATGCTTGTGTCACAAAGATTAA 1099
QY      875 AGATCTTGTGTGATTTTCAAGATGAATAACATTGATGAAAAAGAACTGTTGAATCTTAC 934
Db      1100 AGATCTTGTGTGATTTTCAAGATGAATAACATTGATGAAAAAGAACTGTTGAATCTTAC 1159
QY      935 AAAAGAAAGTCTGCTGATATTTGAAAATTGAAATCTCTAGACTGTGTGTAAGTCT 994
Db      1160 AAAAGAAAGTCTGCTGATATTTGAAAATTGAAATCTCTAGACTGTGTGTAAGTCT 1219
QY      995 GAGGAAAAATTGAAGAGCTCAGACCAAGGTTAAATCCCTTCTCAGGAATTCCTGATGA 1054
Db      1220 GAGGAAAAATTGAAGAGCTCAGACCAAGGTTAAATCCCTTCTCAGGAATTCCTGATGA 1279
QY      1055 ATTATATATGTCATAAATCTAGAGAACTTATGAAAGATCCGGTCATGCGATCAGATGGCTA 1114
Db      1280 ATTATATATGTCATAAATCTAGAGAACTTATGAAAGATCCGGTCATGCGATCAGATGGCTA 1339
QY      1115 TTCTATGAAAAAGAGCAATGGAATAATTTGATCAGAAAAAGAAAGCTCAAGTCCCAT 1174
Db      1340 TTCTATGAAAAAGAGCAATGGAATAATTTGATCAGAAAAAGAAAGCTCAAGTCCCAT 1399
QY      1175 GACAAATCTGTTCTTCTCCTCAGGGTACTTACACCAATAAGACTGTGAATAATGGCCAT 1234
Db      1400 GACAAATCTGTTCTTCTCCTCAGGGTACTTACACCAATAAGACTGTGAATAATGGCCAT 1459
QY      1235 CAATAGATGGCTGGAGACACACCAAAAGTAAA 1266
Db      1460 CAATAGATGGCTGGAGACACACCAAAAGTAAA 1491

RESULT 8
US-10-474-291-40
; Sequence 40, Application US/10474291
; Publication No. US20040132043A1
; GENERAL INFORMATION:
; APPLICANT: AZIMZAI, Yalda; AU-YOUNG, Janice K.
; APPLICANT: BAIKRAI, Sajeev; BAUGHN, Mariah R.
; APPLICANT: BECHA, Shanya D.; BOROMSKY, Mark L.
; APPLICANT: BURFORD, Neil; DING, Li

; APPLICANT: ELLIOTT, Vicki S.; EMERLING, Brooke M.
; APPLICANT: GANDHI, Ameena R.; GIETZEN, Kimberly J.
; APPLICANT: GRIFFIN, Jennifer A.; HAFALIA, April J.A.
; APPLICANT: HONCHELL, Cynthia D.; LAL, Preeti G.
; APPLICANT: LEE, Soo Yeun; LU, Dzung Aina M.
; APPLICANT: ARVIZU, Chandra S.; RAMKOWAR, Jayalaxmi
; APPLICANT: REDDY, Roopa M.; SANJANMALA, Madhusudan M.
; APPLICANT: TANG, Y. Tom; CHAWLA, Narinder K.
; APPLICANT: WANG, Yu-Mei E.; WARREN, Bridget A.
; APPLICANT: XU, Yuming; YANG, Junning
; APPLICANT: YAO, Monique G.; YUE, Henry
; APPLICANT: ZEBARJADIAN, Yeganeh
; TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH
; FILE REFERENCE: PI-0417 USN
; CURRENT APPLICATION NUMBER: US/10/474,291
; CURRENT FILING DATE: 2003-10-06
; PRIOR APPLICATION NUMBER: PCT/US02/11152
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: US 60/349,705
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: US 60/295,263
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/295,340
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/293,727
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/291,846
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/291,662
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: US 60/287,228
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/286,820
; PRIOR FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: US 60/283,294
; PRIOR FILING DATE: 2001-04-11
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 40
; LENGTH: 1773
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 7483131CB1
US-10-474-291-40

Query Match      70.3%; Score 894; DB 18; Length 1773;
Best Local Similarity 80.6%; Pred. No. 4,4e-250;
Matches 1186; Conservative 0; Mismatches 10; Indels 276; Gaps 1;

QY      71 GCGCCGCTCTGACAGCTGTTTCTTCAATAAAGAACATGAGAACTGATTCACAC 130
Db      20 GCGCGGCTGCGGACAGCTGTTTCTTCAATAAAGAACATGAGAACTGATTCACAC 79
QY      131 ATTAGCTGATCATGTGAGAGATGTCAACTGCTGTGCTCTCTTCTTCCCTCTTGCTAC 190
Db      80 ATTAGCTGATCATGTGAGAGATGTCAACTGCTGTGCTCTCTTCTTCCCTCTTGCTAC 139
QY      191 TTGCTCTTGGACAAACAAATTCGCTGACTGTTAGCTGTAAGCTGTAAGTGAAGTGGCACA 250
Db      140 TTGCTCTTGGACAAACAAATTCGCTGACTGTTAGCTGTAAGCTGTAAGTGGCACA 199
QY      251 TTCTCCATGAAGTTTCATACCTATGCTGTGCACTGCTGCTGTTTCCCTTCAGAGACA 310
Db      200 TTCTCCATGAAGTTTCATACCTATGCTGTGCACTGCTGCTGTTTCCCTTCAGAGACA 259
QY      311 TATTTGGCATGTGTTCACACAGATGTACCACTGTCTTATGAGAACTGAAATGGAGACA 370
Db      260 TATTTGGCATGTGTTCACACAGATGTGTACCACTGTCTTATGAGAACTGAAATGGAGACA 319
QY      371 GATGTGGAGATGATGGAACACGCTAAGTGGACGCTGTGAGAGGTTTGGCAATTTTCCCC 430
```

|||||
Db 320 GATCTGGCAGTATGGAACAGCTAGTGGCAGCCCTGTGAGGGTTTGCCAGTTTCCCC 379
Qy 431 AGACTCCAGCTGTTTGGCATCAGGGGCACTGATGGAACCTGGTTTGTGGAATGACCA 490
Db 380 AGACTCCAGCTGTTTGGCATCAGGGGCACTGATGGAACCTGGTTTGTGGAATGACCA 439
Qy 491 GTCATACAAATTAATATGATGTGTAGTGTAAAGATGGCTCCTTGGCGGATGTGCAAT 550
Db 440 GTCATACAAATTAATATGATGTGTAGTGTAAAGATGGCTCCTTGGCGGATGTGCAAT 499
Qy 551 TTCTCCTAATGGAAGCTTCTTGTCTCACTGGCTCCTCACTGGTATTTAAACAGTGGGA 610
Db 500 TTCTCCTAATGGAAGCTTCTTGTCTCACTGGCTCCTCACTGGTATTTAAACAGTGGGA 559
Qy 611 TGATMAAATGAGGTGTCTGCACTAGTGAAGCAAGCATGATCTTGGAAATTAACCTGCTCGA 670
Db 560 TGATMAAATGAGGTGTCTGCACTAGTGAAGCAAGCATGATCTTGGAAATTAACCTGCTCGA 619
Qy 671 TTTTCTTCAAGCAGCTTCTGATGGAAGCAAGCATGATCTTGGAAATTAACCTGCTCGA 730
Db 620 TTTTCTTCAAGCAGCTTCTGATGGAAGCAAGCATGATCTTGGAAATTAACCTGCTCGA 679
Qy 731 ATGTGTCAGGATTTGCCAAGTCAAAATTTGGATTTGTTTCAACCATATCTT----- 784
Db 680 ATGTGTCAGGATTTGCCAAGTCAAAATTTGGATTTGTTTCAACCATATCTTAGGTTT 739
Qy 785 ----- 784
Db 740 TGAATTAATAATATAAAGTACAGTGGGCACTGTGCTCTGCTGTGCTGTGCTT 799
Qy 785 ----- 784
Db 800 TTCCCATGATGGGAGATGCTAGTCTAGGCTCAGTGGATTAAGTCTCATATATATGA 859
Qy 785 ----- 784
Db 860 TACTAATACAGAGATTAATTAATCTTCAACATTTGACTCAGACACAGATATGACAACTTG 919
Qy 785 ----- 784
Db 920 TGCTTTTGACCTAATACCTTTTACTGTCTACTGTCTTAATGGAACAAACAGTGAACAT 979
Qy 785 -----AGCAAGCGGCAAGAAATTCAGCTGGAACA 814
Db 980 CTGGCAATTTGACCTGGAACACTTTGGCCAAAGGACACAAACATCAGCTGAAGCA 1039
Qy 815 ATTTACCGAAGATTTGTCAGAGAGGTCTCTCAACATGGCTTTGTGCACAAGATTTAAA 874
Db 1040 ATTTACCGAAGATTTGTCAGAGAGGTCTCTCAACATGGCTTTGTGCACAAGATTTAAA 1039
Qy 875 AGATCTGTGTGTATTTTCAAGATGAATTAACATTTGATGGAAGAACTGTTGAATCTTAC 934
Db 1100 AGATCTGTGTGTATTTTCAAGATGAATTAACATTTGATGGAAGAACTGTTGAATCTTAC 1159
Qy 935 AAAAGAAAGTGTGCTGATGATTTGAAATTTGAAATCTAGAGTGTGTAAGTCT 994
Db 1160 AAAAGAAAGTGTGCTGATGATTTGAAATTTGAAATCTAGAGTGTGTAAGTCT 1219
Qy 995 GAGGAAATTTGAAGGCTCAGAGCAAGAGTTAAATCCCTTTCTTCAAGAAATCCCTGATGA 1054
Db 1220 GAGGAAATTTGAAGGCTCAGAGCAAGAGTTAAATCCCTTTCTTCAAGAAATCCCTGATGA 1279
Qy 1055 ATTTATATGTCATTAATCTAGAGACTTATGAAAGATCCGCTCATTCGATCAGATGGCTTA 1114
Db 1280 ATTTATATGTCATTAATCTAGAGACTTATGAAAGATCCGCTCATTCGATCAGATGGCTTA 1339
Qy 1115 TTCTATTAAGAAAGAAAGCAATGGAATTTGATGAGAAAGAAAGCAATTAATCCCAT 1174
Db 1340 TTCTATTAAGAAAGAAAGCAATGGAATTTGATGAGAAAGAAAGCAATTAATCCCAT 1399
Qy 1175 GACAAATCTTGTCTTCTCTCAGCGGTACTTACACCAATATGAGACTGGAATATGGCCAT 1234

Db 1400 GACAAATCTTGTCTTCTTCAAGCGGTACTTACACCAATATGAGACTTGAATATGGCCAT 1459
Qy 1235 CAATAGATGGCTGGAGACACCAAAAGTAAA 1266
Db 1460 CAATAGATGGCTGGAGACACCAAAAGTAAA 1491

RESULT 9
US-10-104-047-1689
; Sequence 1689, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1e1 full length cDNA
; FILE REFERENCE: HI-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; PRIOR FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; NUMBER OF SEQ. ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1689
; LENGTH: 1811
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-1689

Query Match 70.0%; Score 890.2; DB 17; Length 1811;
Best Local Similarity 80.9%; Pred. No. 5.8e-249;
Matches 1178; Conservative 0; Mismatches 3; Indels 276; Gaps 1;

Qy 86 GGCCTGTTTCTTCAATTAAGAAACATGGGAAAGTATGATCAGACATTAAGTATGATG 145
Db 125 GGCCTGTTTCTTCAATTAAGAAACATGGGAAAGTATGATCAGACATTAAGTATGATG 184
Qy 146 TGAAGATGTCAGTGTGCTGCTTCTCTTCTCTTCTGCTAATCTTCTTGGACAA 205
Db 185 TGAAGATGTCAGTGTGCTGCTTCTCTTCTCTTCTGCTAATCTTCTTGGACAA 244
Qy 206 AACATTTGCTGTACTCTGTTACGTGATCTTAACTGAACTGCCACATTTCTCAATGAAGTT 265
Db 245 AACATTTGCTGTACTCTGTTACGTGATCTTAACTGAACTGCCACATTTCTCAATGAAGTT 304
Qy 266 TCATACCTAGCTGTGACAGCTGCTGCTTCTCTCTTCCCTTCCAGCATATTTTGGCATGCTG 325
Db 305 TCATACCTAGCTGTGACAGCTGCTGCTTCTCTCTTCCCTTCCAGCATATTTTGGCATGCTG 364
Qy 326 TTCAACAGATGATACCACTGTCTATGGAATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 385
Db 365 TTCAACAGATGATACCACTGTCTATGGAATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 424
Qy 386 GGAACAGCTAGTGGAGCGCCCTGTGAGGGTTTCCAGTTTCCCGACATCTCACGTGTTT 445
Db 425 GGAACAGCTAGTGGAGCGCCCTGTGAGGGTTTCCAGTTTCCCGACATCTCACGTGTTT 484
Qy 446 GGCATCGGGGCGAGCTGATGGAAGTGTGTTTGTGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 505
Db 485 GGCATCGGGGCGAGCTGATGGAAGTGTGTTTGTGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 544
Qy 506 TAGATGTGTAGTGTAAAGTGGCTCTTGGCGGCACTGCAATTTTCTCTAATGAGAG 565
Db 545 TAGATGTGTAGTGTAAAGTGGCTCTTGGCGGCACTGCAATTTTCTCTAATGAGAG 604
Qy 566 CTTCCTTGTCACTGGCTCTCTCATGTGTGATTTAAACAGTGTGGAGATGATTAATATGAGAGT 625
Db 605 CTTCCTTGTCACTGGCTCTCTCATGTGTGATTTAAACAGTGTGGAGATGATTAATATGAGAGT 664
Qy 626 TTGCAATAGTGAAGAAAGCAATATCTGGAATTAACCTGTGGAATTTTCTTCAAGCC 685
Db 665 TTGCAATAGTGAAGAAAGCAATATCTGGAATTAACCTGTGGAATTTTCTTCAAGCC 724
Qy 686 AGTTTGTAGTGAAGAAAGTCTTCAATTTTTCGACTGGACATGATGTGTCAGATG 745

Db 604 GAAAGAGCTTCGATTTTTCGACGTGCATCATGTGCTGAGATTGCCAAGTCAAAATT 663
Qy 759 TGAATGTTCTTTTACCATATCTT----- 784
Db 664 TGAATGTTCTTTTACCATATCTTGTAGTTTGAATTAATAATAAAGTACAGT 723
Qy 785 ----- 784
Db 724 GGGACGTGTCTCTGTCTTGCTTGTGCTTTTCCCATGATGGGAGAGTGTAGTCTCA 783
Qy 785 ----- 784
Db 784 GGGTCACTGATTAAGTCTGTATAGTATAGTACTAATAGTGAATATATCTTACACA 843
Qy 785 ----- 784
Db 844 TTGACTCAGCACACAGGATATGTACAACTTGTGCTTTTGACCTAATACCTTTTACTT 903
Qy 785 ----- 784
Db 904 GCTACTGTTCAATGACAAAGAGTGAACATCTGGCAATTTGACCTGGAAACACTTTC 963
Qy 785 --ACGAAGGCGCAGAAACATCAGTGAAGCAATTTTACCGAAGTTGCTCAGAGAGTCT 842
Db 964 CAACCAAGGCGCAGAAACATCAGTGAAGCAATTTTACCGAAGTTGCTCAGAGAGAT 1023
Qy 843 GTCTCAATGCGCTTGTGCAAGATTTAAAGTCTTGTGATTTTCAAGATGAT 902
Db 1024 GTCTCAATGCGCTTGTGCAAGATTTAAAGTCTTGTGATTTTCAAGATGAT 1083
Qy 903 AACATGATGAAAGAACTGTTGATCTTCAAAAGAAAGTCTGCTGATGATTTGAA 962
Db 1084 AACATGATGAAAGAACTGTTGATCTTCAAAAGAAAGTCTGCTGATGATTTGAA 1143
Qy 963 ATGAATCTTGAAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1022
Db 1144 ATGAATCTTGAAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1203
Qy 1023 GTTAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1082
Db 1204 GTTAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1263
Qy 1083 ATGAAGATCCGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1142
Db 1264 ATGAAGATCCGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1233
Qy 1143 TGAATCAGCAAAAGAAAGTGAAGTCCATGACAAATCTTCTTCTTCTTCTTCTTCTTCTTCTT 1202
Db 1224 TGAATCAGCAAAAGAAAGTGAAGTCCATGACAAATCTTCTTCTTCTTCTTCTTCTTCTTCTT 1383
Qy 1203 CTTCACCAATAGATCTTGAAGTGGCCATCAATAGATGCTGAGAGACACCAAAAG 1262
Db 1384 CTTCACCAATAGATCTTGAAGTGGCCATCAATAGATGCTGAGAGACACCAAAAG 1443
Qy 1263 TAAA 1266
Db 1444 TAAA 1447

RESULT 11
US-10-357-930-29046
; Sequence 29046, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04

PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29046
; LENGTH: 1996
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: misc feature
; LOCATION: 1, 2, 3, 1987, 1988, 1989, 1990, 1991, 1992, 1993, 1994,
; LOCATION: 1995, 1996
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-29046

Query Match 69.1%; Score 878.8; DB 18; Length 1996;
Best Local Similarity 80.7%; Pred. No. 1.3e-245; Indels 276; Gaps 1;
Matches 1166; Conservative 0; Mismatches 2;
Qy 99 CAATAAAGAACATGATGAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 158
Db 4 CAATAAAGAACATGATGAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 63
Qy 159 TGCTGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 218
Db 64 TGCTGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 123
Qy 219 TACTGCTTGAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 278
Db 124 TACTGCTTGAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 183
Qy 279 GTCCAGTGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 338
Db 184 GTCCAGTGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 243
Qy 339 ACCAATGCTTGAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 398
Db 244 ACCAATGCTTGAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 303
Qy 399 GGCAGCCCTGTGAGGTTTGTGCAAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 458
Db 304 GGCAGCCCTGTGAGGTTTGTGCAAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 363
Qy 459 GCTGATGAAGTGTGTTTGTGGAATGACAGTCAATCAATATATATATATATATATATATATATAT 518
Db 364 GCTGATGAAGTGTGTTTGTGGAATGACAGTCAATCAATATATATATATATATATATATATATAT 423
Qy 519 GTTAAGATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 578
Db 424 GTTAAGATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 483
Qy 579 GGTCTCTCATGTGATGATTAACAGTGTGAGATGAATAAATGAGTGTCTGATATGAA 638
Db 484 GGTCTCTCATGTGATGATTAACAGTGTGAGATGAATAAATGAGTGTCTGATATGAA 543
Qy 639 AAAGCAATGATCTTGAATTAACCTGCTGATTTCTTCAAGCAGAGTTCTGATGGA 698
Db 544 AAAGCAATGATCTTGAATTAACCTGCTGATTTCTTCAAGCAGAGTTCTGATGGA 603
Qy 699 GAAAGATCTTCAAGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 758

Db 604 GAACAAGTCTTCAAGTTTTCACATGTCATGTCGATTCAGGATTGCCAAGTCAAAATT 663
QY 759 TGGATTGTTTCTTTTACCATATCTT----- 784
Db 664 TGGATTGTTTCTTTTACCATATCTTGAATTAAATATATAAGTACAGT 723
QY 785 ----- 784
Db 724 GGGCAGTGTCTCTGTTCTGCTTGTCTTTTCCATGATGGCAGATGCTAGTCA 783
QY 785 ----- 784
Db 784 GGGCAGTGTATGCTGTCAATGATATGATATCTAATCTGAGATATATCTTACACA 843
QY 785 ----- 784
Db 844 TTGACTCAGCACAACAGATATGTCACAACCTTGTCTTTTGACCTAATACCTTTACT 903
QY 785 ----- 784
Db 904 GCTACTGTCTCAATGACAAAACAGTGAATCTGCAATTTGACCTGAAACACTTTC 963
QY 785 --AGCAAGGGGCAAGAACTGAGCAATTTTCCGAAGATTGTCAGAGAGTCT 842
Db 964 CAAGCAAGGGGCAAGAACTGAGCAATTTTCCGAAGATTGTCAGAGAGAT 1023
QY 843 GTCTCAACATGAGCTTTGTGCAAGATTTAAAGATCTTGTGATTTTCAAGATGAT 902
Db 1024 GTCTCAACATGAGCTTTGTGCAAGATTTAAAGATCTTGTGATTTTCAAGATGAT 1083
QY 903 AACATGATGAAAAGAACTGTTGATCTTACAAAAGAAAGTCTGCTGATGATTGAAA 962
Db 1084 AACATGATGAAAAGAACTGTTGATCTTACAAAAGAAAGTCTGCTGATGATTGAAA 1143
QY 963 ATGGAATCTCTAGACCTGCGCTAGTAAAGTCTGAGAAATTTGAAGGCTCAGAAC 1022
Db 1144 ATGGAATCTCTAGACCTGCGCTAGTAAAGTCTGAGAAATTTGAAGGCTCAGAAC 1203
QY 1023 GTTAAATCCCTTTCTTCAGGAATCTCTGATGAAATTTATATGTCATATGAGAACTT 1082
Db 1204 GTTAAATCCCTTTCTTCAGGAATCTCTGATGAAATTTATATGTCATATGAGAACTT 1263
QY 1083 ATGAAAGATCCGCTCATGCGATGAGATGCTATTCATATGAAAAGAAAGCAATGAAAAT 1142
Db 1264 ATGAAAGATCCGCTCATGCGATGAGATGCTATTCATATGAAAAGAAAGCAATGAAAAT 1223
QY 1143 TGGATCAGCAAAAAGAAAGTACAGTCCATGACAAATCTTGTCTTCTTCAGCGCTA 1202
Db 1324 TGGATCAGCAAAAAGAAAGTACAGTCCATGACAAATCTTGTCTTCTTCAGCGCTA 1383
QY 1203 CTTACACCAAAATGAGACTCTGAAAATGCGCATCAATGATGGCTGAGACACCAAAAG 1262
Db 1384 CTTACACCAAAATGAGACTCTGAAAATGCGCATCAATGATGGCTGAGACACCAAAAG 1443
QY 1263 TAAA 1266
Db 1444 TAAA 1447

RESULT 12
US-10-077-111-14
; Sequence 14, Application US/10077111
; Publication No. US20020187492A1
; GENERAL INFORMATION:
; APPLICANT: Todderud, C. Gordon
; APPLICANT: Finger, Joshua N.
; APPLICANT: Rillema, J111
; TITLE OF INVENTION: TBA
; FILE REFERENCE: 3053-4114US2
; CURRENT APPLICATION NUMBER: US/10/077,111
; CURRENT FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 60/269,181
; PRIOR FILING DATE: 2001-05-29

QY 12 TCACCTGCGGGGACGTGACCCGACCGCCGCGGACCTTGAAGCGGATCCCGCG 71
Db 37 TCACCTGCGGGGACGTGACCCGACCGCCGCGGACCTTGAAGCGGATCCCGCG 96
QY 72 CCCCCCTCCTGAGGCTGTTTCTTCAAAATPAAAGATGGTGAAGTATTCACACA 131
Db 97 CCCCCCTCCTGAGGCTGTTTCTTCAAAATPAAAGATGGTGAAGTATTCACACA 156
QY 132 TTAGCTGATGATGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 191
Db 157 TTAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 216
QY 192 TGGTCTTGGACAAACAAATGCGCTGATGATGATGATGATGATGATGATGATGAT 251
Db 217 TGGTCTTGGACAAACAAATGCGCTGATGATGATGATGATGATGATGATGATGAT 276
QY 252 TCTCCATTGAAGTTCATATGATGATGATGATGATGATGATGATGATGATGATGAT 311
Db 277 TCTCCATTGAAGTTCATATGATGATGATGATGATGATGATGATGATGATGATGAT 336
QY 337 ATTTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 396
Db 372 ATGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 431
QY 397 ATGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 456
Db 432 GACTCAGCGTGTGGCATCAGGGGAGCGATGATGATGATGATGATGATGATGATGAT 491
QY 457 GACTCAGCGTGTGGCATCAGGGGAGCGATGATGATGATGATGATGATGATGATGAT 516
Db 492 TCATCAAAATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 551
QY 517 TCATCAAAATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 576
Db 552 TCTCTTAATGAAAGCTTCTTGTCACTGCTCTCATGATGATGATGATGATGATGAT 611
QY 577 TCTCTTAATGAAAGCTTCTTGTCACTGCTCTCATGATGATGATGATGATGATGAT 636
Db 612 GATPAAATGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 671
QY 637 GATPAAATGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 696
Db 672 TTTTCTTCAAGCAGCTTCTGATGAGAAAGATGATGATGATGATGATGATGATGATGAT 731
QY 697 TTTTCTTCAAGCAGCTTCTGATGAGAAAGATGATGATGATGATGATGATGATGATGAT 756
Db 732 TGTGTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 784
QY 757 TGTGTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 816
Db 785 ----- 784
QY 817 GAATTAATAATATAAGTACACTGATGAGGACCTGCTCTGTTCTGCTGTGCTTTT 876
Db 785 ----- 784

Db 877 TCCCATGATGGGAGATGCTAGTCTCAGGGTCAGTGATAGTCTGTATATATGAT 936
Qy 785 ----- 784
Db 937 ACTAATAGTGAATTAATTAATCTTCAACATTTGACACACACAGGATATGACAACTTGT 996
Qy 785 ----- 784
Db 997 GCTTTGACCTAATACCTTTTACTTGTCTAGTCTGATGACAAAACAGTGAACATC 1056
Qy 785 -----ACGAAGGCGCAGACGAACATCAGCTGAAGCAA 815
Db 1057 TGGCAATTTGACCTGAAACACATTTGCGACAGGCGCAGACAAACATCAGCTGAAGCAA 1116
Qy 816 TTTTACCAAGATTTGTCAGAGGAGTGTCTCAACATGCTTTGTGTGCAACAGATTTTAAA 875
Db 1117 TTTTACCAAGATTTGTCAGAGGAGTGTCTCAACATGCTTTGTGTGCAACAGATTTTAAA 1176
Qy 876 GATCTGTGTTGATTTTCAAGATGAATTAACATGATGGAAGAAAGACTGTTGAATCTTACA 935
Db 1177 GATCTGTGTTGATTTTCAAGATGAATTAACATGATGGAAGAAAGACTGTTGAATCTTACA 1236
Qy 936 AAAGAAAGTCTGCTGATGATTTGAATAATG----- 966
Db 1237 AAAGAAAGTCTGCTGATGATTTGAATAATGCTGAGATCTCTGCAATGCTCAGCTC 1296
Qy 967 -----AATCTTACGA 977
Db 1297 ACTGCACTCTCAACCTCTGCGGCTCAAGATGATCTCTCACTCTGCAATCTCTTACGA 1356
Qy 978 CTGCGTGTGTAAGTCTGAGGAAATTTGAAGGCTCAGAGCAAGGTTAAATCCCTTCT 1037
Db 1357 CTGCGTGTGTAAGTCTGAGGAAATTTGAAGGCTCAGAGCAAGGTTAAATCCCTTCT 1416
Qy 1038 TCAGAAATTCCTGATGATTTATATATGTCATAAATACTAGAGAACTTATGAAGATCCGCTC 1097
Db 1417 TCAGAAATTCCTGATGATTTATATATGTCATAAATACTAGAGAACTTATGAAGATCCGCTC 1476
Qy 1098 ATGCTATCAGATGCTATTTCTATGTAAGAAAGAGCAATGGAATTTGATGACAAAAG 1157
Db 1477 ATGCTATCAGATGCTATTTCTATGTAAGAAAGAGCAATGGAATTTGATGACAAAAG 1536
Qy 1158 AAAGTCAAGTCCCATGAGCAAAATCTTCTCTCTCAGCGGACTTACCCAAATAGG 1217
Db 1537 AAAGTCAAGTCCCATGAGCAAAATCTTCTCTCTCAGCGGACTTACCCAAATAGG 1596
Qy 1218 ACTCTGAATAATGCGCATCATATAGATGCTGAGACACACCAAAAGTAAA 1266
Db 1597 ACTCTGAATAATGCGCATCATATAGATGCTGAGACACACCAAAAGTAAA 1645

RESULT 13

US-10-077-111-6
; Sequence 6, Application US/10077111
; Publication No. US20020187492A1
; GENERAL INFORMATION:
; APPLICANT: Todderud, C. Gordon
; APPLICANT: Finger, Joshua N.
; APPLICANT: Rillema, Jill
; TITLE OF INVENTION: TBA
; FILE REFERENCE: 3053-6114US2
; CURRENT APPLICATION NUMBER: US/10/077,111
; CURRENT FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 60/294,181
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: 60/269,366
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1901
; TYPE: DNA

ORGANISM: MOUSE
US-10-077-111-6

Query Match 36.5%; Score 464; DB 13; Length 1901;
Best Local Similarity 63.2%; Pred. No. 2.5e-124;
Matches 916; Conservative 0; Mismatches 255; Indels 279; Gaps 2;

Qy 93 TTTCTCAATTAAGAACATGATGTAAGTGAATTCACATTAAGTCTGATGATGATGAT 152
Db 1 TTTCTGTTGTGAGAAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
Qy 153 GTCAATGCTGTGCTTCTCTCTTCTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 212
Db 61 GTCAATGCTGTGCTTCTCTCTTCTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
Qy 213 CGCTGTACTGCTGTACTGATCTTACATGAACTGCAATCTCCATGTAAGTTTCAATCC 272
Db 121 CGCTGTACTGCTGTACTGATCTTACATGAACTGCAATCTCCATGTAAGTTTCAATCC 180
Qy 273 TATGCTGTCACTGCTGCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 332
Db 181 TATGCTGTCACTGCTGCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
Qy 333 GATGTAACAATGCTGTATGAAATTAAGTGAATGACAGATGCTGCAATGATGAAACAG 392
Db 241 GATGTAACAATGCTGTATGAAATTAAGTGAATGACAGATGCTGCAATGATGAAACAG 300
Qy 393 CTAATGTCAGCTGTGATGAGGTTTGCAGTTTCCCAACTTCCAGTGTGTTGGCATCA 452
Db 301 CTAATGTCAGCTGTGATGAGGTTTGCAGTTTCCCAACTTCCAGTGTGTTGGCATCA 360
Qy 453 GGGGAGCTGATGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 512
Db 361 GGGGAGCTGATGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
Qy 513 GGTAGTGTAAAGTGTGCTCTCTGCGGAGTGTGATGATGATGATGATGATGATGATGAT 572
Db 421 GGTAGTGTAAAGTGTGCTCTCTGCGGAGTGTGATGATGATGATGATGATGATGATGAT 480
Qy 573 GTCACTGCTCTCATGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 632
Db 481 GTCACTGCTCTCATGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
Qy 633 AGTGAATAAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 692
Db 541 AGTGAATAAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
Qy 693 GATGAGAACAGTCTTCAATGTTTTCAGTGTGATGATGATGATGATGATGATGATGATG 752
Db 601 GATGAGAACAGTCTTCAATGTTTTCAGTGTGATGATGATGATGATGATGATGATGATG 657
Qy 753 AAAATTTGATGTTTCTTTTACCATATCTT----- 784
Db 658 AAAATTTGATGTTTCTTTTACCATATCTT----- 784
Qy 785 ----- 784
Db 718 CTAATGAGGATGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 777
Qy 785 ----- 784
Db 778 GCATGGGGTCAATGATTAATCTGTATCATATCATATGATGATGATGATGATGATGATG 837
Qy 785 ----- 784
Db 838 CACAGCTGATCAGATACAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 897
Qy 785 ----- 784
Db 898 TTAATGCTACTGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 957
Qy 785 -----ACGAAGGCGCAGACGAACATCAGCTGAAGCAATTTTACGAAGATGATGATG 836

Db 958 CTTGCCAAGCAGAGCATGAAAGCCCTGTAACATTTCACTGAAGATGGTCAG 1017
QY 837 GAGTCGCTCAACATGGCTTTGTGCACAGATTTAAAGATCTGTGGTATTTTCAAG 896
Db 1018 GAGAGATCTCCGGTGGCTTCCGTCTCAAGGCTTGAAGACCTCGTCGGTATTTTCAG 1077
QY 897 ATGAATTAACATTGATGAGAAAGAACTGTGTAATCTTACAAAAGAAAGTGGCTGATGAT 956
Db 1078 GCAACCAACATCATGAGGAAAGAACTATTCATCTCAAAAGGAAAGTCTGGCTGGAT 1137
QY 957 TTGAAATTTGAATCTCTAGACATCGCTAGTAAAGTGTGAGGAAATTTGAAGCTAGG 1016
Db 1138 TTGAAATTCGAATCTCTAGGCTGCGCAGCAAGTCTGTAGAGATTTGAAGAGCTCAGG 1197
QY 1017 ACCAAGGTTAAATCCCTTTCTTCTGAGAAATCTGATGAATTTATATGCTCAATTAACAGA 1076
Db 1198 GCCAAGATGATTCCTCTCTCTTCCGAAATCCCTGACAGATTCATCTGCCAATTAACAGA 1257
QY 1077 GAACTTATGAAAGATCCGCTCATCGCATCAGATGGCTATTCTATATGAAAGAGCAATG 1136
Db 1258 GAACTCATGAAAGACCCGCTCATCGCATCAGATGGCTACTCTTACGAGAGAGAAAGCAATG 1317
QY 1137 GAAATTTGATTCAGCAAAAGAAAGCTGACAAAGTCCCATGACAAATCTTTCTTCTTCA 1196
Db 1318 GAAAGCTGATTCACAAAGAAAGCTGACAGCCCATGACAAATTTGCTCTTCTTCA 1377
QY 1197 GCGGTACTTACACCAATTAAGATCTGAAATGGCCATCAATGATGGCTGAGACAC 1256
Db 1378 CTGTACTTGACCCCAACAGAGACATGAGATGGCCATCAACCGATGGCTGAGACAC 1437
QY 1257 CAAAGTMAA 1266
Db 1438 GAGAGTGAA 1447

RESULT 14
US-10-077-111-5
; Sequence 5, Application US/10077111
; Publication No. US20020187492A1
; GENERAL INFORMATION:
; APPLICANT: Todderud, C. Gordon
; APPLICANT: Finger, Joshua N.
; APPLICANT: Rillema, J111
; TITLE OF INVENTION: TBA
; FILE REFERENCE: 3053-4114US2
; CURRENT APPLICATION NUMBER: US/10/077,111
; CURRENT FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 60/294,181
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: 60/269,366
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 630
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; OTHER INFORMATION: 630 bp partial nucleic acid sequence of human
; OTHER INFORMATION: RET16 cDNA
US-10-077-111-5

Query Match 29.7%; Score 377.2; DB 13; Length 630;
Best Local Similarity 97.9%; Pred. No. 3.1e-99;
Matches 382; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 774 ACCATATCTTAGCAGGCGCAGACATCAGCTGAGCAATTTACGAGATGGTCA 833
Db 241 ACACCTTTCAGCAGGCGCAGACATCAGCTGAGCAATTTACGAGATGGTCA 300
QY 834 GAGAGGTCGTCTCAACATGGCTTTGTGCAACAAGATTTAAAGATCTTGTGTATTTTC 893
Db 301 GAGAGGATGTCTCAACATGGCTTTGTGCAACAAGATTTAAAGATCTTGTGTATTTTC 360

QY 894 AAGATGAATTAACATTGATGAGAAAGAACTGTTGAATCTTACAAAGAAAGTGGCTGAT 953
Db 361 AAGATGAATTAACATTGATGAGAAAGAACTGTTGAATCTTACAAAGAAAGTGGCTGAT 420
QY 954 GATTTGAAATTAATCTCTAGGACTGCGTATGTAAGGTGTGAGGAAATTTGAAGGCTC 1013
Db 421 GATTTGAAATTAATCTCTAGGACTGCGTATGTAAGGTGTGAGGAAATTTGAAGGCTC 480
QY 1014 AGGACCAAGGTTAAATCCCTTTCTTCAAGAAATCTGATGAATTTATATGTCCAATTAAT 1073
Db 481 AGGACCAAGGTTAAATCCCTTTCTTCAAGAAATCTGATGAATTTATATGTCCAATTAAT 540
QY 1074 AGAAGCTTATGAAAGATCCGCTCATCGCATCAGATGAGTGTATTCATATGAAAGAGCA 1133
Db 541 AGAAGCTTATGAAAGATCCGCTCATCGCATCAGATGAGTGTATTCATATGAAAGAGCA 600
QY 1134 ATGAAATTTGATTCAGCAAAAGAAAGCT 1163
Db 601 ATGAAATTTGATTCAGCAAAAGAAAGCT 630

RESULT 15
US-10-357-930-5029/c
; Sequence 5029, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5029
; LENGTH: 446
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-5029

Query Match 26.3%; Score 335; DB 18; Length 446;
Best Local Similarity 96.3%; Pred. No. 5.5e-87;
Matches 365; Conservative 0; Mismatches 10; Indels 4; Gaps 2;
QY 412 GGGTTTCCAGTTTCCCAAGATCCACGTG---TTGGCATAGGGGACGCTGATGAA 468
Db 445 GGTTCACAGTTTTCCTCCAGACTCCCAACGTGTTTGGCATCAGGGGACGCTGATGAA 386
QY 469 CTGTGGTTTGTGAATGACAGCTACATCAAAATTAATATGATGAGTGAAGATG 528
Db 385 CTGTGGTTTGTGAATGACAGCTACATCAAAATTAATATGATGAGTGAAGATG 326
QY 529 GCTCTTGGGCGCATGATTTTCTCTAATGGAAGCTTTGTCACTGAGCTCTCAT 588
Db 325 GCTCTTGGGCGCATGATTTTCTCTAATGGAAGCTTTGTCACTGAGCTCTCAT 266

```

Qy 589 GTGGTGAATTTAACAGTGTGGGATGATAAATGAGGTGCTGCTAGTGTG-AAAAAGCACAT 647
Db 265 GTGGTGAATTTAACAGTGTGGGATGATAAATGAGGTGCTGCTAGTGTG-AAAAAGCACAT 206
Qy 648 GATCTTGAATTTACCTGCTGGGATTTTCTTCACAGCCAGTTTCTGATGAGAACAAAGT 707
Db 205 GATCTTGAATTTACCTGCTGGGATTTTCTTCACAGCCAGTTTCTGATGAGAACAAAGT 146
Qy 708 CTTGAGTTTTTTGCACTGGGCATCATGTGTGTCAGGATTGCCAAGTCAAAATTTGATTGT 767
Db 145 CTTGAGTTTTTTGCACTGGGCATCATGTGTGTCAGGATTGCCAAGTCAAAATTTGATTGT 86
Qy 768 TCTTTACCCATATCTTAG 786
Db 85 TCTTTACCCATATCTTAG 67

```

Search completed: February 5, 2005, 08:35:28
 Job time : 768 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: February 5, 2005, 11:43:17 ; Search time 230 Seconds
(without alignments)
2731.871 Million cell updates/sec

Title: US-10-077-111-13
Perfect score: 2047
Sequence: 1 MVKLHTLADHGDDVNCAP.....LTPNRLTKLAINRLETHOK 384

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues
Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=x1h
-O=/cgn2_1/USPRO.spool/US10077111/runat_04022005_121053_6829/app_query.fasta_1.583
-DB=Issued_Patents_NA -OPMT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPECI=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsu2 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTPMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10077111 @CGN 1.177 @runat_04022005_121053_6829 -NCPU=6 -ICPU=3
-NO MMAP -LARGEOUDRY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued_Patents_NA:
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1978	96.6	1844	4	US-09-620-312D-768
2	234.5	11.5	2359	1	US-08-188-582-4
3	234.5	11.5	2359	1	US-08-188-582-4
4	227	11.1	7028	4	US-09-949-016-4194
5	227	11.1	7028	4	US-09-949-016-4194
6	227	11.1	7042	3	US-09-093-508-1
7	227	11.1	7042	3	US-09-093-508-1
8	227	11.1	7042	3	US-09-093-508-1
9	227	11.1	7042	3	US-09-093-508-1
10	227	11.1	7042	3	US-09-093-508-1
11	227	11.1	7075	3	US-09-093-508-15
12	227	11.1	7075	3	US-09-093-508-15

13	226.5	11.1	2130	4	US-09-620-312D-145	Sequence 145, App
14	220.5	10.8	937	4 <td>US-10-101-464A-251</td> <td>Sequence 251, App</td>	US-10-101-464A-251	Sequence 251, App
15	218	10.6	5152	3 <td>US-09-690-364-10</td> <td>Sequence 10, App</td>	US-09-690-364-10	Sequence 10, App
16	217.5	10.6	3747	3 <td>US-09-690-364-17</td> <td>Sequence 17, App</td>	US-09-690-364-17	Sequence 17, App
17	217.5	10.6	7157	4 <td>US-09-949-016-4192</td> <td>Sequence 4192, App</td>	US-09-949-016-4192	Sequence 4192, App
18	217.5	10.6	7157	4 <td>US-09-949-016-4193</td> <td>Sequence 4193, App</td>	US-09-949-016-4193	Sequence 4193, App
19	217.5	10.6	7171	4 <td>US-09-949-016-971</td> <td>Sequence 971, App</td>	US-09-949-016-971	Sequence 971, App
20	216	10.6	2152	1 <td>US-08-188-582-17</td> <td>Sequence 17, App</td>	US-08-188-582-17	Sequence 17, App
21	215	10.6	2152	1 <td>US-08-188-582-17</td> <td>Sequence 17, App</td>	US-08-188-582-17	Sequence 17, App
22	203.5	9.9	1246	3 <td>US-09-302-769-22</td> <td>Sequence 22, App</td>	US-09-302-769-22	Sequence 22, App
23	198.5	9.7	2166	3 <td>US-09-184-001-1</td> <td>Sequence 1, App</td>	US-09-184-001-1	Sequence 1, App
24	198.5	9.7	2558	3 <td>US-09-184-001-3</td> <td>Sequence 3, App</td>	US-09-184-001-3	Sequence 3, App
25	195	9.5	7886	2 <td>US-08-751-189-2</td> <td>Sequence 2, App</td>	US-08-751-189-2	Sequence 2, App
26	195	9.5	7886	2 <td>US-09-060-836-2</td> <td>Sequence 2, App</td>	US-09-060-836-2	Sequence 2, App
27	195	9.5	7886	2 <td>US-09-184-445-2</td> <td>Sequence 2, App</td>	US-09-184-445-2	Sequence 2, App
28	193.5	9.5	2369	3 <td>US-09-302-769-20</td> <td>Sequence 20, App</td>	US-09-302-769-20	Sequence 20, App
29	188	9.2	3465	3 <td>US-08-914-999-5</td> <td>Sequence 5, App</td>	US-08-914-999-5	Sequence 5, App
30	187.5	9.2	9991	4 <td>US-09-902-540-1014</td> <td>Sequence 1014, App</td>	US-09-902-540-1014	Sequence 1014, App
31	187	9.1	1422	4 <td>US-09-248-796A-4567</td> <td>Sequence 4567, App</td>	US-09-248-796A-4567	Sequence 4567, App
32	184.5	9.0	1157	4 <td>US-09-270-767-14727</td> <td>Sequence 14727, App</td>	US-09-270-767-14727	Sequence 14727, App
33	184.5	9.0	1731	4 <td>US-09-902-540-1920</td> <td>Sequence 1920, App</td>	US-09-902-540-1920	Sequence 1920, App
34	183.5	9.0	1548	4 <td>US-09-614-221A-136</td> <td>Sequence 136, App</td>	US-09-614-221A-136	Sequence 136, App
35	182	8.9	2481	3 <td>US-08-899-578-1</td> <td>Sequence 1, App</td>	US-08-899-578-1	Sequence 1, App
36	180.5	8.8	1638	4 <td>US-09-902-540-8422</td> <td>Sequence 8422, App</td>	US-09-902-540-8422	Sequence 8422, App
37	180.5	8.8	6553	4 <td>US-09-902-540-885</td> <td>Sequence 885, App</td>	US-09-902-540-885	Sequence 885, App
38	180	8.8	1182	4 <td>US-09-248-796A-4568</td> <td>Sequence 4568, App</td>	US-09-248-796A-4568	Sequence 4568, App
39	178.5	8.7	1542	4 <td>US-09-949-016-1365</td> <td>Sequence 1365, App</td>	US-09-949-016-1365	Sequence 1365, App
40	176	8.6	1185	4 <td>US-09-248-796A-4600</td> <td>Sequence 4600, App</td>	US-09-248-796A-4600	Sequence 4600, App
41	176	8.6	2272	3 <td>US-09-108-857-1</td> <td>Sequence 1, App</td>	US-09-108-857-1	Sequence 1, App
42	175.5	8.6	1115	1 <td>US-08-190-802A-19</td> <td>Sequence 19, App</td>	US-08-190-802A-19	Sequence 19, App
43	175.5	8.6	1115	3 <td>US-08-477-346-19</td> <td>Sequence 19, App</td>	US-08-477-346-19	Sequence 19, App
44	175.5	8.6	1115	3 <td>US-08-473-089-19</td> <td>Sequence 19, App</td>	US-08-473-089-19	Sequence 19, App
45	175.5	8.6	1115	3 <td>US-08-487-072A-19</td> <td>Sequence 19, App</td>	US-08-487-072A-19	Sequence 19, App

ALIGNMENTS

RESULT 1
US-09-620-312D-768
Sequence 768 Application US/09620312D
Patent No. 569662
GENERAL INFORMATION:
APPLICANT: Tassat, Tom
APPLICANT: Liu, Chenghua
APPLICANT: Reundi, Vinod
APPLICANT: Zhang, Yie
APPLICANT: Ren, Feiyun
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aiding J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yundong
APPLICANT: Wang, Dunru
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillinghast
APPLICANT: Drmanac, Radojest T.
TITLE OF INVENTION: No. 656962el Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 784CIP28
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pt_Fl_genes Version 1.0
SEQ ID NO 768
LENGTH: 1844
TYPE: DNA


```

REGISTRATION NUMBER: 36, 627
REFERENCE/DOCKET NUMBER: A-57650-2/AUT/RMO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2359 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 49..2160
US-08-188-582-4

Alignment Scores:
Pred. No.: 1,21e-17 Length: 2359
Score: 234.50 Matches: 88
Percent Similarity: 44.34% Conservative: 49
Best Local Similarity: 28.48% Mismatches: 128
Query Match: 11.46% Indels: 45
DB: 1 Gaps: 11

US-10-077-111-13 (1-384) x US-08-188-582-4 (1-2359)
Qy 3 LysLeuIIhEiStHrLeuAlaAspIsgIyAspAspValaEnCyCyAlaAphSer--- 21
Db 1369 GAGGTAACACAGAGAGCTTAATGGGTACACCGGACCCGATACCGGTGGCTTGGCCCC 1422
Qy 22 ---PheSerLeuLeuAlaThrCysSerLeuAspIySthIleArgLeuIySerLeuArg 40
Db 1429 GAGATGAACCTGGTGGCTCATGTTCCGAGGACACACCAATPAGAGCTGTGCTGCTC 1488
Qy 41 AspHethrGluLeuProHisSerProLeuIySPhetiStHrTyAlaValHisCyCyS 60
Db 1489 ACCGTGGCTCGTGA-----GTACCTAACCCGGGACGTTTACCCTGGTGGAGTGT 1542
Qy 61 CysPheSerProSerGIyHisiIleLeuAlaSerCysSerThrAspGIyThrThrValLeu 80
Db 1543 CGCTTGGCGCGCATGGCTCATCTATTGTTCTTCTTTCGTGACGACAAACCTGCTGTG 1602
Qy 81 TrpAsnHrGluAsnGIyGlnMetLeuAlaValMetGluGlnProSerGIySerProVal 100
Db 1603 TGGGCGACGAGATTCACATCAACAGCGTTGGCGGATTCGGGGCACCTG--TCGGACGG 1658
Qy 101 ArgValCysGlnPheSerProAspSerThrCysLeuAlaAlaSerGIyAlaAlaAspGIyThr 120
Db 1660 GATTGTGTACAAATTCATCCCAATTCACATATATGTGGCCACCGGTTCTTAGCGATGCACG 1712
Qy 121 ValValLeuThrPAsnAlaGlnSerTyIyLysLeuTyIyArgCysGlySerValLysAspGIy 140
Db 1720 GTACGCCCTGTGGGACAAACATGACCCGCTGATCGGTACGCTGATGATGAGGGCCACAAGGA 1775
Qy 141 SerLeuAlaAlaCysAlaAphSerProAsnGIySerPhePheValThrGIySerSerCys 160
Db 1780 TCGGTGAGATTCCTGGCCCTTCCGCCGCGCGGTATCTGGGCTCGGGTTCACTAAT 1833
Qy 161 GluAspLeuThrValITrAsp---AspIySHeCArgCysLeuHisSerGIyLysAlaHis 179
Db 1840 CACATCATCATCATCTGGGATCTGTGCAACGGATCCCTGTGACACACCCCTGTGAGGAC 1892
Qy 180 AspLeuGIyIleThrCysCysAspPheSerSerGlnProValSerAspGIyGluGlnGIy 199
Db 1900 ACTAGACACTGTACACACGATCACTTTAGT-----CGGATGGAAACAGTC--- 1944
Qy 200 LeuGlnPhePheArgLeuAlaSerCysGlyGlnAspCysGlnValLysIleITrIleVal 219
Db 1945 -----CTGGCTCGACCGGCTTGATTAACAATCTAATCTGTGG----- 1983
Qy 220 SerPheThrHisIleLeuAlaArgArgThrGluHisGlnLeuLysGlnPheThrGluAsp 239

```

```

Db      1984 -----GACITTCACAAGGTTACCGAAGAC 2007
Oy      240 TrpSerGluGluValIaSerThrTyrLeuCyaaIaGlnAspLeuIysAspLeuValGly 259
Db      2008 TATATTCAGCATCATCATC--ACTGTGTCCGACCATCAGATGAGAAAGC--GAGGAC 2061
Oy      260 IleheIysMetLanAanIleAspGlyIysGlu-----LeuIleAsnLeuThr 275
Db      2062 GTTCACTCACTGCGGTACTCTTCCACGAGAAGAACGTGCCATTGTTCAGCTGCACTTTTACG 2121
Oy      276 LysGluSerLeuAlaAsp-----AspLeuIysIleGluSer 287
Db      2122 CGCCGAATCTCCT-GATGTCCGCGGCTTATTCAAGAGTTAGAGACAGATTAAGCTTA 2180
Oy      288 LeuGlyLeuArgSerIysValLeuArg 296
Db      2181 TTGCTTACTAGCTATGTAAGTTTAAGG 2207

RESULT 3
US-08-646-715-4
Sequence 4, Application US/08646715
Patent No. 5637686
GENERAL INFORMATION:
APPLICANT: Tjian, Robert
APPLICANT: Comati, Lucio
APPLICANT: Dynlact, Brian D.
APPLICANT: Hoey, Timothy
APPLICANT: Ruppert, Siegfried
APPLICANT: Tanese, Naoko
APPLICANT: Wang, Edith
APPLICANT: Weizierl, Robert O.J.
TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAFs AND METHODS OF USE
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESS: FLEHR, HOEBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/646,715
FILING DATE: 09-MAY-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/188,582
FILING DATE: 28-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2359 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 49..2160
US-08-646-715-4

```



```
Qy 163 LeuThrVal1TrpAspArgCysLeuHisSer----- 175
    :::::|||||
    3131 ATCAGGCTCTGGAGACAAAGAAAGTATGTAGAACTGTGTATGTTAAAGCAAGAA 3190
Qy 175 ----- 175
    3191 GTAGATGTGTGTTTCAAGAAATGAAGTGTGCTTGTGACATTGACATATAGAGCT 3250
Qy 176 -----GluTyrAlaHisAspLeuGly1LeuThr 184
    3251 CTGCAACTCATTTATGGAAGACAGTCACTGATTTATCTGACTGAAGCTCAAGTTAGC 3310
Qy 185 CysAspAspPheSerSerGlnProValSerAspGlyGlnGlyLeuGlnPhePheArg 204
    3311 TGCGTTGCTTAAGTCAACAT-----CTTCAGTAC----- 3340
Qy 205 LeuAlaSerCysGlyGlnAspCysGlnValLysIleTrpIleValSerPheThrHisIle 224
    3341 ATTGCATTGTGAGATGAATAATGAGCCATTGAGATTTAGAACTGTAAACAATAGATC 3400
Qy 225 LeuAlaArgArgThrGlnHisGlnLeuLys-----GlnPheThrGluAspTrp 240
    3401 TTCCAGTCCAGGTTTTCAGACAGAAAGAACTGTATGACATCCAGTCCAGCCGAT-- 3457
Qy 241 SerGlnGluValValSerThrTrpLeuCysAlaGlnAsp-LeuLysAspLeuValGly1 260
    3458 GAGAAAGCTCTTATTTCAAGT-----TCTGATGATGTGTAATTCAGATGGAAT 3508
Qy 260 ePheLysMetAsnAsnIleAspGlyLysGlnLeuLysLeuThrLysGlnSerLeuAl 280
    3509 TGGCAATTGGACAATGATGATCTTCTACGAGCCATCAGAAACAGTCAAGACTTAA-- 3566
Qy 280 AlaAspAspLeuLysIleGlnSerLeuGlyLeu-----ArgSerLys 293
    3567 -GACTCTTGAATAATTCAGAGCTGCTTCTGTGATTTGATGAGACAGTGAAGTATGC 3625
Qy 293 eValLeuArgLysIleGlnGluLeuArgThrLysValLysSerLeuSer----- 309
    3626 AATATTAA-----TTACTGGAATAAAGAAAAAACAATTGCTCTCCACAGGATACATA 3679
Qy 310 -----SerGlyIleProAspGluPheIleCysProIleThrArg1 323
    3680 CTTTCTTGTGACATTTCTCAGATGCTACCA--AGTTTCATCTACCTGCTGACAG 3736
Qy 323 uLeuMetLysAspProValIleAlaSer 332
    3737 ACTGCAAGATCTGGAAGTTTGTATCTCC 3764
Db
RESULT 5
US-09-949-016-4195
; Sequence 4195, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4195
; LENGTH: 7028
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4195
```

Alignment Scores:

Pred. No.:	6,63e-16	Length:	7028
Score:	227.00	Matches:	102
Percent Similarity:	40.24%	Conservative:	63
Best Local Similarity:	24.88%	Mismatches:	140
Query Match:	11.09%	Indels:	105
DB:	4	Gaps:	19

US-10-077-111-13 (1-384) x US-09-949-016-4195 (1-7028)

```
Qy 3 LysLeuIleHisThrLeuAlaSerHisGlyAspAspValLysCysValAlaPheSerPhe 22
    2612 GAACAGTACACCTATGATGAGCACTCAGACCAATGCTTCCATTTCCACCAAC 2671
Qy 23 Ser-----LeuAlaIlePheCysSerLeuAspLysThrIleAspLeuTyrSer 38
    2672 AGTAGTCATGATCTTCTTTCAGCAGCTGGGTCAAGTACGCTTCCCAACCTTGGGAT 2731
Qy 39 LeuArgAspPheThrGlnLeuProHisSerProLeuLysPheHisThrTyrAlaValHis 58
    2732 TTGAATGAA-----AAGAAATGTCGAAATACCATGTTGTGATCAATAATTCAGTCAAT 2785
Qy 59 CysCysCysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrThr 78
    2786 CACTGACGATTTTTCACCAAGATGATTAAGCTTTGGCTAGTTGTTACGCTGAGAACCTTA 2845
Qy 79 ValLeuTrpAsnThrGlnGlnGly-----GlnMetLeu 89
    2846 AAGCTTTGGATGCGCATGACAGAAATGAGAGAAAGCAATTAATGTGAACAGTTCTTC 2905
Qy 90 AlaValMetGlnGlnProSerGlySerPro-----ValArgValCysGlnPheSer 106
    2906 CTAAATTTGAGACACCTCAAGAGATATGAAATGATATGTAAGCTTGTGCTGCTCT 2965
Qy 107 ProAspSerThrCysLeuAlaSerGlyAlaAlaAspGlyThrValValLeuTrpAsnAla 126
    2966 GCTGATGCTGCAAGATATATGATG--GCAGCAAAAATAAATCTTTTGTGGAAATACA 3022
Qy 127 GlnSer---TyrLysLeuTyrArgCysGlySerValLysAspGlySerLeuAla----- 143
    3023 GACTCAGGTTCAAGAGGCTGATGCTG--AGAGACACTTTTAAGTGGGTT 3070
Qy 144 ---AlaCysAlaPheSerProAsnGlySerPhePheValThrGlySerSerCysGlyAsp 162
    3071 CATGCTGATGATGTTTCTCTGATGATCATCATTTTGTACATCTTGTGATGACCAACA 3130
Qy 163 LeuThrVal1TrpAspArgCysLeuHisSer----- 175
    3131 ATCAGGCTCTGGAGACAAAGAAAGTATGTAGAACTGTGTATGTTAAAGCAAGAA 3190
Qy 175 ----- 175
    3191 GTAGATGTGTGTTTCAAGAAATGAAGTGTGCTTGTGACATTGACATATAGAGCT 3250
Qy 176 -----GluTyrAlaHisAspLeuGly1LeuThr 184
    3251 CTGCAACTCATTTATGGAAGACAGTCACTGATTTATCTGACTGAAGCTCAAGTTAGC 3310
Qy 185 CysAspAspPheSerSerGlnProValSerAspGlyGlnGlyLeuGlnPhePheArg 204
    3311 TGCGTTGCTTAAGTCAACAT-----CTTCAGTAC----- 3340
Qy 205 LeuAlaSerCysGlyGlnAspCysGlnValLysIleTrpIleValSerPheThrHisIle 224
    3341 ATTGCATTGTGAGATGAATAATGAGCCATTGAGATTTAGAACTGTAAACAATAGATC 3400
Qy 225 LeuAlaArgArgThrGlnHisGlnLeuLys-----GlnPheThrGluAspTrp 240
    3401 TTCCAGTCCAGGTTTTCAGACAGAAAGAACTGTATGACATCCAGTCCAGCCGAT-- 3457
Qy 241 SerGlnGluValValSerThrTrpLeuCysAlaGlnAsp-LeuLysAspLeuValGly1 260
    3458 GAGAAAGCTCTTATTTCAAGT-----TCTGATGATGCTGAATTCAGGATGGAAT 3508
Db
```



```

Db      3509 TGGCAATTGACAAATATCTTTCTACGAGGCCATCAGAAACATGAAAGACTTAA-- 3566
Qy      280 aAaPpLeuLysIleGluSerLeuGlyLeu-----ArgSerly 293
Db      3567 -GACTCTTGAATAAATTCAGACCTGCTTTCTTGTCATTTGATGGAACAGTGAAGTATGC 3625
Qy      293 sValLeuArgLysIleGluGluLeuArgThrLysValLysSerLeuSer----- 309
Db      3626 AATATT-----TTACTGGAATAAAGAAAAGACCTTGCTGTCTACACAGGTTACAGTA 3679
Qy      310 -----SerGlyIleProAspGluPheIleCysProIleThrArgI 323
Db      3680 CTTTCTTGACATTTCTACAGATGCTACCA---AGTTTCATCTACCTTGCTGACAG 3736
Qy      323 uLeuMetLysAspProValIleAlaSer 332
Db      3737 ACTGCAAAAGATCTGAGATTTTGATCTCC 3764

RESULT 7
US-09-435-115-1
; Sequence 1, Application US/09435115
; Patent No. 6346607
; GENERAL INFORMATION:
; APPLICANT: Henzel, William J.
; TITLE OF INVENTION: APAF-1, AN ACTIVATOR OF C ASPASE-3
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
; STREET: 3100 No. 634607west Center, 90 South Seventh St
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/435,115
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/092,508
; FILING DATE:
; APPLICATION NUMBER: 60/055,258
; FILING DATE: 07-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Kettelberger, Ph.D., Denise M
; REGISTRATION NUMBER: 33,924
; REFERENCE/DOCKET NUMBER: 11669.6USU1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-332-5300
; TELEFAX: 612-332-9081
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7042 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 578...4159
; OTHER INFORMATION:
; US-09-435-115-1

Alignment Scores:
Pred. No.: 6.65e-16 Length: 7042
Score: 227.00 Matches: 102
Percent Similarity: 40.24% Conservative: 63

```

```

Best Local Similarity: 24.88% Mismatches: 140
Query Match: 11.09% Indels: 105
DB: 3 Gaps: 19

US-10-077-111-13 (1-384) x US-09-435-115-1 (1-7042)
Qy      3 LysLeuLleIleIleThrLeuAlaAspHisGlyAspArgValAlaGlyCysAlaPheSerPhe 22
Db      2612 GAAGTGTACACACCTATGATGAGCATCTAGAGCAAGTCAATTCCTGTCATTCACCAAC 2671
Qy      23 Ser-----LeuLeuAlaThrCysSerLeuAspLysThrIleArgLeuTyrSer 38
Db      2672 AGTAGTCATCATCTTCTCTTAGCCACCTGGCTCAAGTCACTGCTCCCAAACTTTGGGAT 2731
Qy      39 LeuArgAspPheThrGluLeuProHisSerProLeuLysPheHisThrTyrAlaValHis 58
Db      2732 TTGAATCAA-----AAAGAAATGCGAAATACCAATGTTTGCTCATACAAATTCAGTCAAT 2785
Qy      59 CysCysAspPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrThr 78
Db      2786 CACTGCAGATTTTTCACCAAGATGATTAAGCTTTTGCTTAGTTGTTCAAGCTGATGGAACCTTA 2845
Qy      79 ValLeuTrpAsnThrGluAsnGly-----GlnMetLeu 89
Db      2846 AAGCTTTGGGATGCGACATCAGCAAAATGAGAGAAACATTAATGTAAGCAAGTTCTTC 2905
Qy      90 AlaValMetGluGlnProSerGlySerPro-----ValArgValCysGlnPheSer 106
Db      2906 CTAAATTTGAGAGACCTCAAGAGATGAGAAAGTATGATGAAGTGTGTTCGTGGCT 2965
Qy      107 ProAspSerThrCysLeuAlaSerGlyAlaAlaAspGlyThrValValLeuTrpAsnAla 126
Db      2966 GCTGATGTCGCAAGATTAATGCTG---GCAAGCAAAATAAATCTTTTGTGGAATACA 3022
Qy      127 GlnSer---TyrLysLeuTyrArgCysGlySerValLysAspGlySerLeuAla----- 143
Db      3023 GACTCAGCTTCAAAAGGCGCTGATGC-----AGAGACCATTTTAAGTTGGGCT 3070
Qy      144 ---AlaCysAlaPheSerProAsnGlySerPheValThrGlySerSerCysGlyAsp 162
Db      3071 CATGTGTGATGTTTCTCTGATGATCATTTTTCACATCTTGTGATGACACAGACA 3130
Qy      163 LeuThrValTrpAspAspLysMetArgCysLeuHisSer----- 175
Db      3131 ATCAGGCTCTGGAGACAAAGAAAGATGTAAGAACTCTCTGTAATGTTAAAGCAAGAA 3190
Qy      175 ----- 175
Db      3191 GTAGATGTGTGTTTCAAGAAATGAAGTATGATGCTCTTGACATGACATTAATAGACGT 3250
Qy      176 -----GlnLysAlaHisAspLeuGlyIleThr 184
Db      3251 CTGCAACTCATTAATGGAAGAACAGGTCAATGATTAATGACTGAAGCTCAAGTTAGC 3310
Qy      185 CysCysAspPheSerSerGlnProValSerAspGlyGluGlnGlyLeuGlnPhePheArg 204
Db      3311 TGCTGTGCTTAAGTCACAT-----CTTCAGTAC----- 3340
Qy      205 LeuAlaSerCysGlyGlnAspCysGlnValLysIleTrpIleValSerPheThrHisIle 224
Db      3341 ATTCGATTTGAGATGAATAATGAGGCATTAAGATTTTAAGACTGTGTAACCAATAGATC 3400
Qy      225 LeuAlaArgArgThrGluHisGlnLeuLys-----GlnPheThrGluAspTrp 240
Db      3401 TTCAGTCCAGGTTTCAGCAAGAAACTGTATGCGACATCCAGTTCACAGCCGAA--- 3457
Qy      241 SerGluGluValValSerThrTrpLeuGlyAlaGlnAsp-LeuLysAspLeuValGlyI 260
Db      3458 GAGAAGACTTTATTTTCAGT-----TTGATGATGCTGAATTCAGGATGGAAT 3508
Qy      260 ePheLysMetAsnAsnIleAspGlyLysGluLeuAsnLeuThrLysGluSerLeuAl 280
Db      3509 TGGCAATTGACAAATATCTTTCTACGAGGCCATCAGAAACATGAAAGACTTAA-- 3566

```

```

OY 280 aAspAspLeuLySIleGluSerLeuGlyLeu-----ArgSerLy 293
Db 3567 -GACCTTGAATAAAATTCAGACTGCTTCTTGTCATTGATGGAACAGTAGATAGC 3625
OY 293 sValLeuArgLySIleGluLeuArgThrIleValLySerLeuSer----- 309
Db 3626 AATATTA-----TTACTGGAATAATAAAAAAGACTTGTCTGTGCACAGGTACAGTA 3679
OY 310 -----SerGlyIleProAspGluPheIleCysProIleThrArgI 323
Db 3680 CTTTCTTGACATTCTTCAGATGCTACCA---AGTTTCATCTACTCTGTGACAG 3726
OY 323 uLeuMetLyAspProValIleAlaSer 332
Db 3737 ACTGCMAAGATCTGGAGTTTGATCTCC 3764

RESULT 8
US-09-098-310-1
; Sequence 1, Application US/09098310
; Patent No. 6403765
; GENERAL INFORMATION:
; APPLICANT: Alnemuri, Emad S.
; TITLE OF INVENTION: TRUNCATED APAF-1 AND METHODS OF USE
; FILE REFERENCE: 480140.438
; CURRENT APPLICATION NUMBER: US/09/098,310
; CURRENT FILING DATE: 1998-06-16
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 7042
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (578) ... (4159)
US-09-098-310-1

Alignment Scores:
Pred. No.: 6,656-16 Length: 7042
Score: 227.00 Matches: 102
Percent Similarity: 40.24% Conservative: 63
Best Local Similarity: 24.88% Mismatches: 140
Query Match: 11.09% Indels: 105
DB: 3 Gaps: 19

US-10-077-111-13 (1-384) x US-09-098-310-1 (1-7042)
OY 3 LysIleuIleHisThrLeuAlaAspHisGlyAspValIaenCysValAPheserPhe 22
Db 2612 GAACATAGTACACACTTATGATGACGACATCAGACGAAATGCAATGTGTCATTTCACCAAC 2671
OY 23 Ser-----LeuLeuAlaThrCysSerLeuAspLyThrIleArgLeuTySer 38
Db 2672 AGTACTCATCATCTTCTCTAGCCCTGAGCTCAAGTCACTGCTCTCCAAACTTGGGAT 2731
OY 39 LeuArgAspPheThrGluLeuProHisSerProLeuLySPheHisThrTyAlaValHis 58
Db 2732 TTGAATCAA-----AAAGAAATGTCGAATATCCATGTTGGTCACAAATTCAGTCAT 2785
OY 59 CysCysCysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrThr 78
Db 2786 CACTGCAGATTTTTCACAGATGATATAGCTTTGGCTAGTGTTCACGTGATGGAACCTTA 2845
OY 79 ValIleTPAenThrGluAsnGly-----GlnMetLeu 89
Db 2846 AAGCTTTGGGAGCCACATACAGCAATGAGAGAAACATTAATGTGAACAGTTCTTC 2905
OY 90 AlaValMetGluGlnProSerGlySerPro-----ValArgValCysGlnPheSer 106
Db 2906 CTAATTTGGAGACCTTCMAAGAGATATGGAAGTATGAGATGTTGGTTGCGTGCT 2965

```

QY	107	ProbaSerThrCysLeuAlaSerGlyAlaAlaAspGlyThrValValLeuThrPheAla	126
Db	2366	GCTATGGTGCACGATATATGTCG--GCAGCAAAAATAAAATCTTTTGGGAATACA	3022
QY	127	GlnSer---TyrIysLeuTyrArgCysGlySerValIysAspGlySerLeuAla-----	143
Db	3023	GACTCAGCTTCAAAGGTCGCTGATTCG-----AGAGCAACTTTTAAGTTGGCGTT	3070
QY	144	---AlaCysAlaPheSerProAsnGlySerPhePheValThrGlySerSerCysGlyAsp	162
Db	3071	CATGGTGTCGATGCTTTCTCTCGATGATCATCATTTTGTGACATCTTCTGATGCACAGACA	3130
QY	163	LeuThrValIlePheAspArgIysMetArgCysLeuHisSer-----	175
Db	3131	ATCAGGCTCTGGAGACAAAGAAAGTATGTGAAGAACTCTGTAATGTTAAAGCAAGA	3190
QY	175	-----	175
Db	3191	GTAATCTGTGCTTTCAAGAAATGAAGTATGCTCTTCGACTGACATATATAGACGT	3250
QY	176	-----GlnIysAlaHisAspLeuGlyIleThr	184
Db	3251	CTGCACACTCATTATGGAAGACAGATCAGATTGATATGCTGACTGAAGCTCAAGTTAGC	3310
QY	185	CysCysAspPheSerSerGlnProValSerAspGlyGlnGlyLeuGlnPhePheArg	204
Db	3311	TGCTGTGGCTTTAACTCCACAT-----CTTCAGTAC-----	3340
QY	205	LeuAlaSerCysGlyGlnAspCysGlnValIysIleThrIleValSerPheThrHisIle	224
Db	3341	ATTGCATTTGGAGATGAAATGAGCCATTGACATTTTAAGAACTGTGTAACATAGATC	3400
QY	225	LeuAlaArgArgThrGlnHisGlnLeuLys-----GlnPheThrGlnAspTrp	240
Db	3401	TTCACAGTCCAGGTTTACAGACAAAGAAATCTGATGACATCCAGTCCAGCCGAT---	3457
QY	241	SerGlnGlnValIysSerThrTrpLeuCysAlaGlnAsp-LeuLysAspLeuValGlyIle	260
Db	3458	GAGAAAGACTTATATTTCAAGT-----TCTGATATCTGAAATTCAGATGGAAT	3508
QY	260	ePheLysMetAsnAsnIleAspGlyIysGlnLeuLeuAsnLeuThrIysGlnSerLeuAl	280
Db	3509	TGGCAATTTGGACAAATATGATCTTTCTACAGGCGCATAGAAACATGAAAGACTTAA--	3566
QY	280	AspAspLeuLysIleGlnSerLeuGlyLeu-----ArgSerIle	293
Db	3567	-GACTCTGAAATAATTCACAGCTGCTTCTTGTCATTGATGGAACAGTGAAGGTATGG	3625
QY	293	SValLeuArgIysIleGlnIleuLeuArgThrIysValIysSerLeuSer-----	309
Db	3626	AATATTAA-----TTACTGGAAATAAGAAAGAACTTTGCTGTCAACAGGCTACAGTA	3679
QY	310	-----SerGlyIleProAspGlnPheIleCysArgProIleThrArgGly	323
Db	3680	CTTTCCTTGTCACATTTCTACGATGCATCAACA---AGTTTCATCTACTCTGCTGACAG	3736
QY	323	uLeuMetLysAspProValIleAlaSer	332
Db	3737	ACTGCAAGATCTGAGATTTTGAATCTCC	3764
RESULT 9			
US-09-690-364-21			
; Sequence 21, Application US/09690364			
; Patent No. 6468795			
; GENERAL INFORMATION:			
; APPLICANT: Hong Zhang			
; APPLICANT: Andrew T. Walt			
; TITLE OF INVENTION: ANTISENSE MODULATION OF APPA-1 EXPRESSION			
; FILE REFERENCE: RTS-0190			
; CURRENT APPLICATION NUMBER: US/09/690,364			
; CURRENT FILING DATE: 2000-10-17			
; NUMBER OF SEQ ID NOS: 100			
; SEQ ID NO 21			

```

; LENGTH: 7042
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (578) ... (41622)
US-09-690-364-21

```

Alignment Scores:	
Pred. No.:	6.65e-16
Score:	237.00
Percent Similarity:	40.24%
Best Local Similarity:	24.88%
Query Match:	11.09%
DB:	3
Length:	704
Matches:	102
Conservative:	63
Mismatches:	140
Indels:	105
Gaps:	19

US-10-077-111-13 (1-384) X US-09-690-364-21 (1-70421)

Qy	3	LysLeuIleHsStrLeuAlaAspHisGlyAspValIleAsnCySAlAspSerHe	22
Db	2612	GAACtAGACACACcCTATGATGAGACACTCAGACAGCAATGCTGCGCATTTACCAAC	2671
Qy	23	Se-----LeuLeuAlaThrCySserLeuAspLysrHleArgLeuTyrSer	38
Db	2672	AGTATGCATCATCTTCTCTTAGCCCACTGGGTCAAGTGACTGCTCTTCAAACTTTGGGAT	2733
Qy	39	LeuArgAspRheThrGluLeuProHisSerProLeuYsPheHisThrTyrAlaValHis	58
Db	2732	TTGAATCA-----AAAGAAATGTCGAATACCAATGTTTGGTCATACAAATTCACATCAT	2785
Qy	59	CysCySAspPheSerProSerGlyHisIleLeuAlaSerCySserThrAspGlyThrThr	78
Db	2786	CACtGCAGATTTTACCAcGATGATTAAGCTTTTGGCTATGTTTACAGCTGATGGAACCTTA	2844
Qy	79	ValLeuTyrPasnThrGluAsnGly-----GlnMetLeu	89
Db	2846	AAGCTTTGGGATGGCAGCATCAGCAAAATGAGAGAAAACATTAATGTGMAACAGTTCTTC	2905
Qy	90	AlaValMetGluGlnProSerGlySerPro-----ValArgValCySglnPheSer	106
Db	2906	CTAAATTTGGAGAACCTTCAAGAGAGATATGGAAGTATAGTGAAGATGTTGTCGGGACT	2965
Qy	107	ProAspSerThrCySLeuAlaSerGlyValIleAlaAspGlyThrValValLeuTyrPasnIle	126
Db	2966	GCTATGCTGTCAGAGATTAATGCTG--CCAGCAAAAAATAAATCTTTTGTGGAAATCA	3022
Qy	127	GlnSer---TyrLysLeuTyrArgCySgLySerValLysAspGlySerLeuAla-----	143
Db	3023	GACTCACGCTTCAAAAGTGGCTGATTTGC-----AGAGCACTTAAAGTTAGTGGGTT	3070
Qy	144	---AlaCySAlaPheSerProAsnGlySerPhePheValIThrGlySerSerCySgLyAsp	162
Db	3071	CATGGCTGATGATGTTTCTCTGATGATGATCATCATTTTGGACATCTTGATGATGACCAACA	3130
Qy	163	LeuThrValTyrAspAspLysIleMetArgCySLeuHisSer-----	175
Db	3131	ATCAGGCTCTGGGAGACAAABAAGTAATGTAAGAACTCTGCTGTATATGTTAAAGCAAAA	3199
Qy	175	-----	175
Db	3191	GTAGATGTGTGTTTCAAGAAATGAACTGATGCTCTTGACAGTTGACCATATTAAGAGCT	3250
Qy	176	-----GluLysAlaHisAspLeuGlyLLeuThr	184
Db	3251	CTGCACATCTTAATATGGAAGAACGCTCAGATTGTTATCTGACTGAAAGCTCAACATTAGC	3310
Qy	185	CysCySAspPheSerSerGlnProValSerAspGlyGluGlnGlyLeuGlnPhePheArg	204
Db	3311	TGCTGTTCTTAAGTCCACAT-----CTTCAGATAC-----	3340
Qy	205	LeuAlaSerCySgLyGlnAspCySglnValLysIleTrrPleValSerPheThrHisIle	224
Db	3341	ATTGCGATTGGATGAAATGAAAGCCCATGTGATTTTGAACCTGTAAACATATGAATC	3400

Qy	225	neulIaArgxgThrGlnHIEgIUleuLys-----GlnPhetHrGlnsApr	240
Db	3401	TTCCAGTCCAGCGTTTTCACACAGAAACTGATGGACATCCAGTTTCACAGCCGAT---	3457
Qy	241	SerGlnGluValValSerThrTrpLeuGlyCysAlaGlnAsp-LeuLysAspLeuValGlyI	260
Db	3458	GAGAAAGACTCTTATATTTCAMGT-----TCTGATGATGCTGAATTCAGGATGGAAT	3508
Qy	260	ePhelysMetCysAsnIleLeaGlyLysGluLeuLysLeuIleThrLysGlnSerLeuAl	280
Db	3509	TGGCAATTGGACAAATGATGATCTTTCTACAGAGCCGATCAGAAAACAGTGAAGACITTTA--	3566
Qy	280	AspAspLeuLysIleGlnSerLeuGlyLeu-----ArgSerLys	293
Db	3567	-GACTCTTGAAAATAATTCAGAGCTGTTCTTGTCATATTGATGGAACAGTGAAAGGATATGG	3625
Qy	293	sValLeuArgLysIleGlnGluLeuArgThrLysValLysSerLeuSer-----	309
Db	3626	AAATATTA-----TTACTGGAATAATGAAAAAGACTTGTGTCGACACAGGTAACAGTA	3679
Qy	310	-----SerGlyIleProAspGluPheIleCysProIleThrArgG	323
Db	3680	CTTTCCTTGTCACATTTCTCAGAGTCCTACCA---AGTTTCATCTCACTCTGCTGCACAAAG	3736
Qy	323	uLeuMetLysAspProValIleLeaSer	332
Db	3737	ACTGCAAAAGATCTGGACGTTTGTGATCTCC	3764

RESULT 10
100 00 040 010 100

```

US-09-949-016-159
; Sequence 159, Application US/09949016
; Patent No. 6812339
;
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;
; FILE REFERENCE: CU001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
;
; NUMBER OF SEQ ID NOS: 207012
;
; SOFTWARE: FastSeq for Windows Version 4.0

```

```

; SEQ ID NO 159
; LENGTH: 7042
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-159

Alignment Scores:
Pred. No.:      6,656-16
Score:          227.00
Percent Similarity: 40.24%
Best local Similarity: 24.88%
Query Match:    11.09%
DB:              4
                Length: 7042
                Matches: 102
                Conservative: 63
                Mismatches: 140
                Indels: 105
                Gaps: 19

```

US-10-077-111-13 (1-384) X US-09-949-016-159 (1-7042)

[illegible]

[illegible]

```

DB      3737 ACTGCAAGATCTGGAGTTTGATCTCC 3764

RESULT 11
US-09-092-508-15
; Sequence 15, Application US/09092508
; Patent No. 6291643
GENERAL INFORMATION:
APPLICANT: Henzel, William J.
TITLE OF INVENTION: APAF-1, AN ACTIVATOR OF C ASPASE-3
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
STREET: 3100 No. 6291643west Center, 90 South Seventh St
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FaesSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/092,508
FILING DATE: 05-JUN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/048,807
FILING DATE: 05-JUN-1997
APPLICATION NUMBER: 60/055,258
FILING DATE: 07-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: Kettelberger, Ph.D., Denise M
REGISTRATION NUMBER: 33,924
REFERENCE/DOCKET NUMBER: 11669.6USU1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-332-5300
TELEFAX: 612-332-9081
TELEX:
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 7075 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 578..4192
OTHER INFORMATION:
US-09-092-508-15

Alignment Scores:
Pred. No.:          6.7e-16           Length:       7075
Score:              227.00            Matches:       102
Percent Similarity: 40..24%          Conservative: 63
Best Local Similarity: 24.88%        Mismatches:   140
Query Match:        11.09%           Indels:       105
DB:                  3                Gaps:         19

US-10-077-111-13 (1-384) x US-09-092-508-15 (1-7075)
QY      3 LysLeulleIhstHrLwAlAaPhIaGLyAsPaRvAlAaNCySvAlApheSerPhe 22
      :::::::::::::::::::: : : : : : : : : : : : : : : : : : : : : : :
Db      2645 GAACTAGACACACTTGTGATGAACACTTCAGAGCAAGTCAATTGCTCCATTTCACCAC 2704
QY      23 Ser-----LauLeuAlaThrCySeSerLeuAbpLyStHrIlleArgLeuTySer 38
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      2705 AGTAGCATCATCTTCTCTTAGCACCATGGGTCAGAGTGCATCGCTTCCCAAACTTGGGAT 2764
QY      39 LeuArGaRpHeThnrgLulePurProHisSerProLeuLyBheHtIstHyTlAlaValHis 58
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
      :::: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      2765 TTGATGCAA-----AAAGAATGTGCAAAATGCCATGTTGGTGCATACMAATTTCAGTCAT 2818

```

OY	59	CysCyCyPheSerProSerGlyHisIleLeuIleAspGlyThrThr	78
DB	2819	CACCTGCAGATTTTACCCAGATGATTAAGCTTTGGCTAAGTTTCAAGCTGATGAACTTA	2878
OY	79	ValLeuTrpAenThrGluAsnGly-----GlnMetLeu	89
DB	2879	AAGCTTTGGATGGCACTTCAGCAAAATGAGAGAAACATTAAATGTGAACAGTCTTC	2938
OY	90	AlaValMetGluGlnProSerGlySerPro-----ValArgValCysGlnPheSer	106
DB	2939	CTAAATTTGGAGGACCCTCAAGAGGATATGAAAGTAGATAGTGAAGTGTGTTGGTGGCT	2998
OY	107	ProApeSerThrCysLeuAlaSerGlyValAlaAspGlyThrValValLeuTrpAsnAla	126
DB	2999	GCTGATGGTGCAGAGATTAATAGTGTG---CCAGCAAAAATAAATCTTTTGTGGAAATCA	3055
OY	127	GlnSer---TyrLysLeuTyrArgCysGlySerValLysAspGlySerLeuAla----	143
DB	3056	GACTCACGCTTCAAAAGGTGGCTGATGTGC-----AGAGCATTATTAAGTGGGCT	3103
OY	144	---AlaCysAlaPheSerProAsnGlySerPhePheValThrGlySerSerCysGlyLysP	162
DB	3104	CATGGTGTGATGATGTTTCTCTGATGATGATCAATTTTGGACATCTTCGATGACCAACA	3163
OY	163	LeuThrValTyrPaspAspLysMetArgCysLeuHisSer-----	175
DB	3164	ATCAGCGCTCTGGAGACAAAGAAAGATATGTAAAGACTCTGCTGTATATGTTAAAGCAAGA	3223
OY	175	-----	175
DB	3224	GTAAGATGTGTTTCAGAAAATGAAAGTAGTGCTTCGACGTTGACATATTAAGCGT	3283
OY	176	-----GluLysAlaHisAspLeuGlyIleThr	184
DB	3284	CTGCACATCATTTAATGAAAGAACAGGTCAAGTATGATTCGACTGAAGCTCAAGTTAGC	3343
OY	185	CysCysAspPheSerSerGlnProValSerAspGlyGlnGlyLeuGlnPhePheArg	204
DB	3344	TGCTGTGCTTAAAGTCCACAT-----CTTCAGATAC-----	3373
OY	205	LeuAlaSerCysGlyGlnAspCysGlnValLysIleTyrPheIleValSerPheThrHisIle	224
DB	3374	ATTGATATTGGAGATGAATAAATGAGCCCATGGAATTTAAGACTTGAACAATATGAATC	3433
OY	225	LeuAlaArgTyrThrGluHisGlnLeuLys-----GlnPheThrGluAspTrp	240
DB	3434	TTCCAGCTCCAGGTTTCACACACAGAAAATCTGATGCGCATCCACAGCTTACAGCGCAT---	3490
OY	241	SerGlnGluValValSerThrTyrLeuCysAlaGlnAsp-LeuLysAspLeuValGlyIle	260
DB	3491	GAGAAAGACTCTTATATTTCAAGT-----TCTGATGATGCTGAATTCAGATATGAGAT	3541
OY	260	ePheLysMetAsnAsnIleAspGlyLysGlnLeuLeuAsnLeuThrLysGlnSerLeuAl	280
DB	3542	TGGCAATTGGACAAATATATCTTTCTACAGAGCCCATCAGAAACAGTGAAGAACTTTA-	3599
OY	280	AspAspLeuLysIleGlnSerLeuGlyLeu-----ArgSerLys	293
DB	3600	-GACTCTTGAATAAATTCAAGACTGCTTCTTGTCATTTGATGAAACAGTAAGGTATGG	3658
OY	293	eValLeuArgLysIleGlnGluLeuArgThrLysValLysSerLeuSer-----	309
DB	3659	AATATTA-----TTACTGGAATATAAAGAAAAGCTTGTTGTCACACAGGATACAGTA	3712
OY	310	-----SerGlyIleProAspGlnPheHisCysProIleThrArgI	323
DB	3713	CTTCTTGATGACATTTTCACAGATGCTACCA---AGTTTATATCACTCGTGCAGAAAG	3766
OY	323	uLeuMetLysAspProValIleAlaSer	332
DB	3770	ACTGCAAGATCTGAGTTTATCTCC	3797

```

RESULT 12
US-09-435-115-15
; Sequence.15, Application US/09435115
; Patent No. 634607
; GENERAL INFORMATION:
; APPLICANT: Henzel, William J.
; TITLE OF INVENTION: APAF-1, AN ACTIVATOR OF CASPASE-3
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
; STREET: 3100 No. 634607west Center, 90 South Seventh St
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/435.115
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/092,508
; FILING DATE:
; APPLICATION NUMBER: 60/055,258
; FILING DATE: 07-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ketelberger, Ph.D., Denise M
; REGISTRATION NUMBER: 33,924
; REFERENCE/DOCKET NUMBER: 11669.6USUI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-332-5300
; TELEFAX: 612-332-9081
; TELEX:
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7075 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 578...4192
; OTHER INFORMATION:
; US-09-435-115-15

Alignment Scores:
Pred. No.: 6.7e-16 Length: 7075
Score: 227.00 Matches: 102
Percent Similarity: 40.24% Conservative: 63
Best Local Similarity: 24.88% Mismatches: 140
Query Match: 11.09% Indels: 105
DB: 3 Gaps: 19

US-10-077-111-13 (1-384) x US-09-435-115-15 (1-7075)

QY 3 LysLeuLeuLeuIsthrLeuAlaAspHisGlyAspAspValaenCyGValaPheSerPhe 22
Db 2645 GAACCTAGTACACACCTTATGATGACACCTCAGACGACGAAGTCAATTCTGTCATTTACCAAC 2704
QY 23 Ser-----LeuLeuAlaThrCysSerLeuAspLysThrIleArgLeuTyrSer 38
Db 2705 AGTGGTATCATCTTCTCTTGGCCACTGGGCTCAAGTACAGCTGCTTCTCCAAACCTTGGGAT 2764
QY 39 LeuArgAspPheThrGluLeuProHisSerProLeuLysPheHisTyrAlaValHis 58
Db 2765 TTGAATCA-----AAAGAATGTGCAAAATACCATGTTTGGTCTATACAAATTCAGTCAAT 2818
QY 59 CysCysCysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrThr 78

```

```

Db      2819 CACTGCAGATTTCACCGATGATGATGACCTTTGGCTAGTGTTCAGCTGATGGAACCTTA 2878
Qy      79  ValLeuTriphenThrgluangly-----GlmLeu 89
Db      2879 AAGCTTTGGATCGCATCAGCAATGAGAGAAAAGCATTAATGGAACAGTCTTC 2938
Qy      90  AlaValMeGlunInProSerGlySerPro-----ValArgValCysGlnPheSer 106
Db      2939 CTAATTTGGAGGACCTCAAGAGATATGGAAGTATGGAAGTGTGTTGCTGTGCT 2998
Qy      107 ProAspSerThrCysLeuAlaSerGlyAlaAlaAspGlyThrValValLeuTriphenAla 126
Db      2999 GCTGATGGTCGCAAGATMATGTG--GCAGCAAAAATAAATCTTTTGTGGAAATACA 3055
Qy      127 GlnSer---TyrIysLeuTyrArgCysGlySerValIysAspGlySerLeuAla----- 143
Db      3056 GACTCAGCTTCAAAGGCGCTGATTGC-----AGAGACATTTAAAGTTGGGTT 3103
Qy      144 ---AlaCysAlaPheSerProAsnGlySerPhePheValThrGlySerSerCysGlyAsp 162
Db      3104 CATGTGTGATGTCTTCTCTGATGATCATCTTTTGACATCTTCTGATGACAGACA 3163
Qy      163 LeuThrValITrpaAspIlyPheAspIysCysLeuHisSer----- 175
Db      3164 ATCAGGCTCTGGAGACAAAGAAAGTATGTAGAAACTCTGTATATGTTAAAGCAAGAA 3223
Qy      175 ----- 175
Db      3224 GTAGATGTTGTGTTTCAAGAAAATGAAAGTATGTCCTTGACATGACCATATMAAGCT 3283
Qy      176 -----GlnIysAlaHisAspLeuGlyIleThr 184
Db      3284 CTCGAATCATTAATGAGAAACAGGTACAGATTGATTATCTGACTGAAGCTCAAGTTAC 3343
Qy      185 CysCysAspPheSerSerGlnProValSerAspGlyGlnGlnGlnPhePheArg 204
Db      3344 TCGTGTGCTTAAGTCCACAT-----CTTCAGTAC----- 3373
Qy      205 LeuAlaSerCysGlyGlnAspCysGlnValIleTrpIleValSerPheThrHisIle 224
Db      3374 ATTGCATTTGGAGATGAAAATGAGCCATTGAGATTTTAAACAATTGAAACAATAGACAT 3433
Qy      225 LeuAlaArgThrGlnHisGlnLeuIys-----GlnPheThrGlnAspTrp 240
Db      3434 TTCACATCCAGGTTTCAGCAAGAAAGCTGATGACATCCAGTTCAACAGCCGAT--- 3490
Qy      241 SerGlnGluValIleSerThrTrpLeuCysAlaGlnAsp-LeuIysAspLeuValGlyI 260
Db      3491 GAGAAAGCTCTTAATTCAGT-----TCTGATGATGCTGAAAATTCAAGTATGGAAT 3541
Qy      260 ePheIysMetAsnAsnIleAspGlyIysGlnLeuLeuAsnLeuThrIysGlnSerLeuAl 280
Db      3542 TGGCAATTGGACAAAGTATCTTTCTACAGGCCATCAGAAACAGTGAAGACTTTA-- 3599
Qy      280 aaAspAspLeuIysIleGlnSerLeuGlyLeu-----ArgSerIly 293
Db      3600 -GACTCTGAAAATTCACAGACTCTTCTTGTGATCTTATGATGAACAGTGAAGGATATG 3658
Qy      293 sValIleuArgIleIleGlnGlnLeuArgThrIysValIysSerLeuSer----- 309
Db      3659 AATATTA-----TTACTGGAATAAAGAAAAGACTTTGTCTGTCCACAGGCTACAGTA 3712
Qy      310 -----SerGlyIleProAspGlnPheIleCysProIleThrArgI 323
Db      3713 CTTTCTTGAGACATTTCTACGATGCTACA---AGTTTCATCTACTGCTGACAG 3769
Qy      323 uIeuMeClyAspProValIleAlaSer 332
Db      3770 ACTGCAAAAGATCTGAGATTTCATCTCC 3797

```

```

/ Sequence 145, Application US/09620312D
/ Patent No. 6569662
/ GENERAL INFORMATION:
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Zhang, Jie
/ APPLICANT: Ren, Feiyan
/ APPLICANT: Chen, Rui-hong
/ APPLICANT: Zhao, Qing A.
/ APPLICANT: Wehrman, Tom
/ APPLICANT: Xue, Aidong J.
/ APPLICANT: Yang, Yonghong
/ APPLICANT: Wang, Jian-Rui
/ APPLICANT: Zhou, Ping
/ APPLICANT: Ma, Yundong
/ APPLICANT: Wang, Dunru
/ APPLICANT: Wang, Zhwei
/ APPLICANT: John Tillinghaet
/ APPLICANT: Drmanac, Radoje T.
/ TITLE OF INVENTION: No. 656962e1 Nucleic Acids and
/ FILE REFERENCE: 784C1P2B
/ CURRENT APPLICATION NUMBER: US/09/620,312D
/ CURRENT FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: 09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: 09/488,725
/ PRIOR FILING DATE: 2000-01-21
/ NUMBER OF SEQ ID NOS: 1105
/ SOFTWARE: pc_fl_genes Version 1.0
/ SEQ ID NO 145
/ LENGTH: 2130
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (184)..(1236)
/ US-09-620-312D-145

Alignment Scores:
Pred. No.: 9,6e-17 Length: 2130
Score: 226.50 Matches: 96
Percent Similarity: 41.05% Conservative: 60
Best Local Similarity: 25.26% Mismatches: 137
Query Match: 11.06% Indels: 88
DB: 4 Gaps: 15

US-10-077-111-13 (1-384) x US-09-620-312D-145 (1-2130)
Qy      25  LeuAlaThrCysSerLeuAspIlySerThrIleArgLeuTyrSerLeuArgAspPheThrGlu 44
Db      149  CTAGCTACAGCTTTCTGAAGACAAATCATTAATAGTATGAGCATG-----TATGCCGAG 202
Qy      45  LeuProHisSerProLeuIysPheHisThrTyrAlaValHisCysCysCysAspPheSerPro 64
Db      203  CGCTCCCTGATATCTTGATGACATACACACTGTGACGCTGCGGCCCA-ATTGTAACC 261
Qy      65  SerGlyHisIleLeuAlaSerCysSerThrAspGlyThrValLeuTriphenThrGlu 84
Db      262  GATGGAAGCTAATGTGTGATGTAGTGAATAAACTATTAAATTTGGGATACACACA 321
Qy      85  AsnGlyGlnMeLeuAlaValIleMetGlnInProSerGlySerProValArgValCysGln 104
Db      322  AATAGCAATGTGTTAATATACCTTCAGATTCGCTGGATTT--GCAATTTTGTGAC 378
Qy      105  PheSerProAspSerThrCysLeuAlaSerGlyAlaAlaAspGlyThrValValLeuTrp 124
Db      379  TTAAACCTAGAGGTATCATGATGATCTTCAGACGCTTCGATCAAACTGGAAGCTGG 438
Qy      125  AsnAlaGlnSerTyrIysLeuTyrArgCysGlySerValIysAspGlySerLeuAlaI 144
Db      439  GATGTAAAGTGAACAATTAATTAATCAACGATTAACAAGTTCAACAGCGGTGAGATTATTC 498

```


Db 716 GTGCCCCATTTCCTGTCGCCCATCTTCAGAAATAATGATGATCTTCAGATTGCG 775
QY 332 SerAspGlyTyrSerTyrGlyValMetGluAsnTrpIleSerIleValArg 351
:::|||||:::|
Db 776 GCGGATGGTTTCACTTATGAAAGGAGATGCGGGAATGGTTGGACAACGACGAT 835
QY 352 ThrSerProMetThrAsnLeuValLeuProSerAlaValLeuThrProAsnAlaGlyTrp 371
|||:::|||||
Db 836 ACATCCCAATGACCAACTTGAAAGCTTATGATTCATTCATCTCACCACGCGGTTA 895
QY 372 LysMetAlaIleAsnArgTrpLeu 379
:::|||||
Db 896 CGGTCCGCAATTCAGAGTGGCTC 919
RESULT 15
US-09-690-364-10
; Sequence 10. Application US/09690364
; Patent No. 6468795
; GENERAL INFORMATION:
; APPLICANT: Hong Zhang
; APPLICANT: Andrew T. Walt
; TITLE OF INVENTION: ANTISENSE MODULATION OF APAF-1 EXPRESSION
; FILE REFERENCE: RTS-0190
; CURRENT APPLICATION NUMBER: US/09/690,364
; CURRENT FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 100
; SEQ ID NO 10
; LENGTH: 5152
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (586)...(4302)
US-09-690-364-10
Alignment Scores:
Pred. No.: 4,83e-15 Length: 5152
Score: 218.00 Matches: 65
Percent Similarity: 39.15% Conservative: 27
Best Local Similarity: 27.66% Mismatches: 75
Query Match: 10.65% Indels: 68
DB: 3 Gaps: 6
US-10-077-111-13 (1-384) x US-09-690-364-10 (1-5152)
QY 3 LysLeuIleHisThrLeuAlaAspHisGlyAspAspValAsnCysCysAlaPheSerPhe 22
|||||:::|
Db 2620 AACCTTGTGCACACCTTACGACGACGACCTGGAGCAAGTCAATTGCTGCCATTTCACCAAC 2679
QY 23 Ser-----LeuLeuAlaThrCysSerLeuAspLysThrIleArgLeuTyrSer 38
|||:::|||||
Db 2680 AGTAGTAACCACTTCTTGGCCACCGGGTCAATGATTTCTTCTCAAGCTCTGGGAT 2739
QY 39 LeuArgAspPheThrGluLeuProHisSerProLeuLysPheHisThrTyrAlaValHis 58
|||:::|||||
Db 2740 TTGAATCAA-----AAGAAATGTCGAAATACCATGTTGGTGTGTCACAGAACTCAGTCAAC 2793
QY 59 CysCysCysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrTrp 78
|||:::|||||
Db 2794 CACTGCAAGTTCTCACCAGACGATGAGCTTTGGCTAGCTGCTCAGCTGACGGGACTTTA 2853
QY 79 ValLeuTrpAsnThrGluAsnGlyGln----- 87
|||:::|
Db 2854 AGGCTTTGGGATGTGAGATCAGCAAAAGAGAGAAAGATTAATGTAAGCGCTTCTTC 2913
QY 87 ----- 87
Db 2914 CTGAGTTCAGAAAGACCTCCAGAGATGTGAGATGATCTGTGAAGTGTGTTCTGTGTC 2973
QY 87 ----- 87
Db 2974 GCAGATGTGCAAAATATATAGTGGACGAAAAAACAAGTCTCTTTTGTGATTTCA 3033

QY 88 -----MetLeuAlaValMetGluGlnProSerGlySerProValArgValCysGln 104
:::|||||:::|
Db 3034 ACTAGTGCCCTATTGGCAGAGATCCACAGGCCATTCACGACCATTCAGTACTGTGAC 3093
QY 105 PheSerProAspSerThrCysLeuAlaSerGlyAlaAlaAspGlyThrValValLeuTrp 124
|||||:::|
Db 3094 TTCTCCCTTAGACCATTTGGCTGTGATGGCTGTCTGACGACTGTGTGAGATTGTGG 3153
QY 125 AsnAlaGlnSer---TyrIleValTyrArgCysGlySerValLysAspGlySerLeuAla 143
|||:::|
Db 3154 AACATAGACTCCCGCCTAAGGTGGCGGACTGC-----AGAGACATTTGAGT 3201
QY 144 -----AlaCysAlaPheSerProAsnGlySerPhePheValThrGlySerSerCys 160
|||||:::|
Db 3202 TGGGTTACAGGTGTGTGATGTTTCTCCCGATGGCTCCTCATTTTGTGACAGCTTCTGATGAC 3261
QY 161 GlyAspLeuThrValThrAspArgLysMetArgCysLeuHisSer 175
:::|||||
Db 3262 CAACATAGAGGTCTGGGAGACAAATAAGTATGCAAGACTCT 3306

Search completed: February 5, 2005, 14:56:56
Job time : 263 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 5, 2005, 08:18:36 ; Search time 4596 Seconds
(without alignments)
10534.759 Million cell updates/sec

Title: US-10-077-111-12

Perfect score: 1272

Sequence: 1 gaatcgcgttcctcctcg.....acaccaaagtaagaattc 1272

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 3966416

Minimum DB seq length: 0
Maximum DB seq length: 200

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST.*
1: gb_esc1.*
2: gb_esc2.*
3: gb_hnc.*
4: gb_esc3.*
5: gb_esc4.*
6: gb_esc5.*
7: gb_esc6.*
8: gb_gss1.*
9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	199	15.6	199	2	BE669639 7e16e07.x
C 2	199	15.6	199	2	BR110272 7n51g07.x
C 3	156.8	12.3	160	1	AA501581 ne97g11.s
C 4	130.2	10.2	199	1	AA1415375 mc35f10.x
C 5	111.8	8.8	190	2	BF653223 276923.MA
C 6	80.8	6.4	138	6	CA587431 LBEL2034
C 7	78.8	6.2	82	1	AA912881 c127b06.s
C 8	71.4	5.6	191	2	BB323974 BB323974
C 9	49.2	3.6	177	2	BB186455 BB186455
C 10	45.4	3.6	102	7	CR455364 CR455364
C 11	42.8	3.4	189	2	AM032239 EST725693
C 12	39.8	3.1	184	9	AG211342 Oryza sat
C 13	39.2	3.1	144	1	AT622018 486031E09
C 14	34.2	2.7	175	2	BE240075 BE240075
C 15	33.8	2.7	195	7	CF370509 IG50B08.Y
C 16	33.6	2.6	172	1	AF071906 AF071906
C 17	33.6	2.6	172	1	AF071907 AF071907
C 18	33.4	2.6	165	7	CP968712 CP968712
C 19	33.2	2.6	168	9	CL790258 OR_BB010
C 20	33.2	2.6	168	9	CL790258 OR_BB010
C 21	33.2	2.6	179	1	AV174594 AV174594
C 22	33	2.6	181	7	N89830 ZB34G04.s1
C 23	33	2.6	183	6	C01352 HMG5000833
C 24	32.6	2.6	130	8	BZ394951 E1NAJ35TR

C 25	32.4	2.5	121	2	AM039819
C 26	32.4	2.5	149	1	AF109011
C 27	32.2	2.5	137	1	AV289619
C 28	31.8	2.5	170	2	AZ721313
C 29	31.8	2.5	171	2	AM022465
C 30	31.8	2.5	171	2	AM022465
C 31	31.6	2.5	158	1	AA577442
C 32	31.6	2.5	170	7	CF526046
C 33	31.2	2.5	130	5	BP519881
C 34	31.2	2.5	148	7	W49852
C 35	31.2	2.5	152	2	BF361608
C 36	31.2	2.5	156	6	CD989424
C 37	31.2	2.5	159	7	CV296852
C 38	31.2	2.5	189	9	CR044488
C 39	31	2.4	151	5	BP699577
C 40	31	2.4	158	5	BM154677
C 41	31	2.4	178	7	CV427795
C 42	30.8	2.4	200	1	AU089064
C 43	30.8	2.4	171	4	BG266397
C 44	30.8	2.4	186	4	BG272274
C 45	30.6	2.4	146	7	CF249442

ALIGNMENTS

RESULT 1
LOCUS BE669639/c
DEFINITION 7e16e07.x1 NCI CGAP Lu24 Homo sapiens CDNA clone IMAGE:328260 3' similar to TR:080828 080828 HYPOTHETICAL 88.8 KD PROTEIN. ;, mRNA sequence.

ACCESSION BE669639
VERSION BE669639.1
KEYWORDS GI:10030180
SOURCE Homo sapiens
ORGANISM Homo sapiens (human)

REFERENCE
AUTHORS NCI-CGAP
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lemon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LINL, send email to: info@image.lnl.gov

Trace considered overall poor quality
Seg primer: -40up from Gibco
High quality sequence stop: 1.
Location/Qualifiers
1. 199
/organism="Homo sapiens"
/mol_type="RNA"
/db_xref="taxon:9606"
/clone="IMAGE:328260"
/tissue_type="carcinoid"
/lab_host="DH10B"
/note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI-CGAP-Lus was prepared, and ss circles were used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clonoids

FEATURES

source

ORIGIN

1414920-1417991 and 1520904-1522439). Subtraction by Bento Soares and M. Fatima Bonaldo. "

Query Match 15.6%; Score 199; DB 2; Length 199;

Best Local Similarity 100.0%; Pred. No. 1.5e-45;

Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN

Soares and M. Fatima Bonaldo. "

Query Match 15.6%; Score 199; DB 2; Length 199;

Best Local Similarity 100.0%; Pred. No. 1.5e-45;

Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1056 TTTATATGTCATTAACCTAGAGAACTTATGAAAGATCCGTCATCGCATGAGTGGCTAT 1115
DB 199 TTTATATGTCATTAACCTAGAGAACTTATGAAAGATCCGTCATCGCATGAGTGGCTAT 140
QY 1116 TCATATGAAAAGAGACATGAAAATTGGATCAGCAAAAAGAAAGCTACAGTCCCATG 1175
DB 139 TCATATGAAAAGAGACATGAAAATTGGATCAGCAAAAAGAAAGCTACAGTCCCATG 80
QY 1176 ACAATCTGTCTTCTTCAGCGGTACTACACCAATAGSACTCGAAATGGCCATC 1235
DB 79 ACAATCTGTCTTCTTCAGCGGTACTACACCAATAGSACTCGAAATGGCCATC 20
QY 1236 AATAGATGCTGGAGACAC 1254
DB 19 AATAGATGCTGGAGACAC 1

RESULT 2

BF110272 199 bp mRNA linear EST 20-OCT-2000

LOCUS 7n51g07.x1 NCI CGAP Lu24 Homo sapiens CDNA clone IMAGE:3568309.3,
DEFINITION similar to TR:Q9SYV6 Q9SYV6 HYPOTHETICAL 48.9 KD PROTEIN.;, mRNA
sequence.

ACCESSION BF110272 GI:10939962
VERSION BF110272
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 199)
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
COMMENT Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov

Trace considered overall poor quality
Seq primer: -40UP from Gibco
High quality sequence stop: 1.
Location/Qualifiers
1. 199
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3568309"
/issue_type="carcinoid"
/lab_host="DH10B"
/clone_1lb="NCI-CGAP Lu24"
/note="Organ: lung; Vector: pTT3D-Pac (Pharmacia) with a
modified polylinker; Plasmid DNA from the normalized
library NCI CGAP Lu24 was prepared, and ss circles were
made in vitro. Following HAP purification, this DNA was
used as tracer in a subtractive hybridization reaction.
The driver was PCR-amplified cDNAs from a pool of 5,000
clones made from the same library (cloneids
1414920-1417991 and 1520904-1522439). Subtraction by Bento

FEATURES

source

QY 1044 ATTCTGATGAATTTATATGTCCTCAATTAAGAGAACTTATGAAAGATCCGTCATCGCA 1103
DB 199 ATTCTGATGAATTTATATGTCCTCAATTAAGAGAACTTATGAAAGATCCGTCATCGCA 140
QY 1104 TCAGATGCGTATTCATATGAAAAGAGCAATGCAAAAATTGGATCAGCAAAAAGAAAGCT 1163
DB 139 TCAGATGCGTATTCATATGAAAAGAGCAATGCAAAAATTGGATCAGCAAAAAGAAAGCT 80
QY 1164 ACAATGCCATGACCAATCTGTCTTCTTCAGCGGTACTTACACCAATAGSACTCTG 1223
DB 79 ACAATGCCATGACCAATCTGTCTTCTTCAGCGGTACTTACACCAATAGSACTCTG 20
QY 1224 AAAATGCCATCAATAGAT 1242
DB 19 AAAATGCCATCAATAGAT 1

RESULT 3

AA501581 160 bp mRNA linear EST 19-AUG-1997

LOCUS ne97g11.81 NCI CGAP Kid1 Homo sapiens CDNA clone IMAGE:912260
DEFINITION similar to SW:PKM THECU P49695 PUTATIVE SERINE/THREONINE-PROTEIN
KINASE PKM.;, mRNA sequence.

ACCESSION AA501581 GI:2236548
VERSION AA501581
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 160)
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
COMMENT Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: David B. Kitzman, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bdrp/image/image.html

Trace considered overall poor quality
Insert Length: 519 Std Error: 0.00
Seq primer: -40m3 fwd. ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1. 160
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:912260"
/issue_type="kidney"
/lab_host="DH10B"
/clone_1lb="NCI-CGAP Kid1"
/note="Vector: pAMP10; mRNA made from invasive kidney
tumor, cDNA made by oligo-dT priming. Non-directionally
cloned. Size-selected on agarose gel, average insert
size 600 bp. Reference: Krizman et al. (1996) Cancer
Research 56:5380-5383."

FEATURES

source

ORIGIN

	Query Match	12.3%	Score 156.8	DB 1	Length 160
	Best Local Similarity	98.8%	Pred. No. 1.8e-33		
	Matches 158	Conservative 0	Mismatches 2	Indels 0	Gaps 0
Qy	249	CATTCTCATGAAAGTTTCATACCTATGCTGTCCACTGCTCTGTTCTCCCTTCAGGA	308		
Db	160	CATTCTCATGAAAGTTTCATACCTATGCTGTCCACTGCTCTGTTATCCCTTCAGGA	101		
Qy	309	CATATTTTGGCATGTGTTCACACAGATGTACCACTGTCTTATGGAAATACTGAAAATGGA	368		
Db	100	CATATTTTGGCATGTGTTCACACAGATGTACCACTGTCTTATGGAAATACTGAAAATGGA	41		
Qy	369	CAGATGCTGGCAGTATGTAAGCAGCTATGTGGCAGCCCTG	408		
Db	40	CAGATGCTGGCAGTATGTAAGCAGCTATGTGGCAGCCCTG	1		

RESULT 4	LOCUS	DEFINITION
AI415375/c	AI415375	199 bp mRNA linear EST 09-FEB-1999
	mc55f10.x1	Soares mouse pJNMF19.5 Mus musculus cDNA clone

ACCESSION	AI415375	
VERSION	AI415375.1	GI:4258879
KEYWORDS	EST.	
SOURCE	Mus musculus	(house mo
ORGANISM	Mus musculus	

REFERENCE
AUTHORS

TITLE	The WashU-NCI Mouse EST Project 1999
JOURNAL	Unpublished (1999)
COMMENT	Contact: Marra M/WashU-NCI Mouse EST Project 1999

Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
This clone was previously sequenced on the 5' end only, this new data is from the 3' end
Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
High quality sequence stop: 1.
Location/Qualifiers
1..199

```

/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:350539"
/dev_stage="19.5 dpc total fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_1fb="Soares mouse P3MMP19.5"
/notes="vector: pT7T3D (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer (5',
TGTTACCAATCTGAAGTGGAGACGGCCATTTTTTTTTTTTTTTT 3'),
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M. Fatima Bonaldo. RNA was kindly provided by
Dr. Minoru Ko (Wayne State University)."
```

	Query Match	10.2% ; Score 130.2 ; DB 1 ; Length 199 ;
	Best Local Similarity	78.4% ; Prod. No. 8,6e-26 ;
	Matches 156; Conservative	0; Mismatches 43; Indels 0; Gaps 0;
QY	1044	ATTCTGATGAATTTATATGTCCAAATACTAGAGAACTTATGAAAAGATTCGGTCATCGCA 1103
Db	199	ATCCTTGACCAAGTTCATAGTCACAAATAACCAAGAACTCATGAAAGACCCTGCATCGCA 140
QY	1104	TCAAGATGGCTATTTCATATGAAAAAGAAGCAATGGAATAATTGSATCAGCAAAAAGAAACGT 1163
Db	139	TCAGATGGCTAATCTCTACGAGAGAAGCAATGCAAAAGCTGSAATCCACAAGAAAGAACGT 80
QY	1164	ACAAGTCCCAATGACAAATCTTGTCTTCCTTCAGCGGTACTTACACCAAAATGAGACTCTG 1223
Db	79	ACCGGCCCAATGTCAAATTTGGCGCTTCATTCATCGTAAGTGAACCCAAAACAGACACTG 20
QY	1224	AAAATGGCCATCAATAGAT 1242
Db	19	AAGATGGCCATCAACCAAT 1

RESULT 5	BP653223/c	LOCUS	BF653223	190 bp	mRNA	linear	EST 25-APR-2001
DEFINITION	2769923	MARC 3BOV Bos taurus	CDNA 5', mRNA sequence.				
FEATURES	2769923						

VERSION	BF653223.1	GI:11918292
KEYWORDS	EST.	
SOURCE	Bos taurus	(cow)
ORGANISM	Bos taurus	

REFERENCE
AUTHORS

REFERENCE	1 (Pages 1 to 190)
AUTHORS	Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Cesaes,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-Mckown,C.G., Pertea,G., Holt,I.I., Karamycheva,S., Liang,F., Quackenbush,J. and Keelke,J.W.
TITLE	Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle
JOURNAL	Genome Res. 11 (4), 626-630 (2001)
MEDLINE	21180013
PUBMED	11282978
COMMENT	Contact: Smith TPL

COMMENT

Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4380
Email: smitht@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -mnscore 18
and -minmatch 12 options.

PCR primers
FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTTCCAGTCACGACG
Plate: 68 row: H column: 24
Seq primer: ATTATGGTGACCATATAG.

```

FEATURES
source
location/Qualifiers
1..190
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_1b="MARC 3BOV"
/note="vector: pCMV SPORT6; site 1: NotI; Site 2: SalI;
library made from pooled tissue from marrow, alveolar
macrophage, ovary, fetal semitendonsus muscle, and fetal
longissimus muscle."

```

ORIGIN

Query Match 8.8%; Score 111.8; DB 2; Length 190;

Best Local Similarity 87.4%; Pred. No. 1.6e-20;
Matches 146; Conservative 0; Mismatches 17; Indels 4; Gaps 2;

QY 1104 TCAGATGGCTATTCATATGAGAAAGCAATGAGAAATTCAGCAAAAGAAAGCT 1163

DB 181 TTGAGAGGGCTACTCATATGAGAAAGCAATGAGAAATTCAGCAAAAGAAAGCT 122

QY 1164 ACAAGTCCCATGACAAATCTGTTCTTCCTT-CAGCGGTACTTACACCAATAGACTCT 1222

DB 121 ACCAGTCCCATGACAAATCTGTTCTTCTTACAGTACACTTACTCAATATAGACTCT 62

QY 1223 GAAATATGGCCATCATATGATGGCT--GGAGACACACCAAAAGTAA 1266

DB 61 GCAAAATGGCCATCATATGATGGCTAACTGAGACACATCAAAATGAA 15

RESULT 6
LOCUS CA587431 138 bp mRNA linear EST 12-JAN-2004

DEFINITION LBE12P34 cDNA from mouse aorta Mus musculus cDNA, mRNA sequence.

ACCESSION CA587431 GI:40792674

VERSION CA587431.1

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Borang, S., Andersson, T., Thelin, A., Odberg, J. and Lundberg, J.
1 (bases 1 to 138)
Vascular gene expression in atherosclerotic plaque prone regions
analysed by representational difference analysis

JOURNAL Unpublished (2002)

COMMENT Contact: Andersson Tove
Department of Biotechnology

TEKNIKTINGEN 34, plan 6, 100 44 Stockholm, Sweden

Tel: +46 8 790 71 29

Fax: +46 8 245452

Email: tove@biochem.kth.se

Representations (amplified cDNA) from plaque prone regions

Seq primer: CTA TGA CCA TGA TTA CGC CAA G.

Location/Qualifiers

1. .138

/organism="Mus musculus"

/mol_type="mRNA"

/strain="ADP/- and LPLR/- on C57BL/6x 129 background"

/db_xref="taxon:10090"

/sex="male"

/dev_stage="8 weeks old"

/clone_1lb="cDNA from mouse aorta"

/note="Organ: aorta; Site 1: DpnII; Site 2: DpnII; cDNA

was prepared from whole aorta divided in atherosclerotic

plaque prone regions (aortic arch and abdominal aorta

proximal part) and less plaque prone regions (descending

thoracic aorta and abdominal aorta distal part). cDNA was

fragmented with DpnII, linker ligated and amplified to

generated starting material for representational
difference analysis (RDA). The two cDNA pools were
subjected to iterative RDA subtraction and amplification
to enrich for gene fragments differentially expressed at
early stages of atherosclerosis."

ORIGIN

Query Match 6.4%; Score 80.8; DB 6; Length 138;

Best Local Similarity 81.0%; Pred. NO. 1.1e-11;

Matches 94; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 1151 CAAAAAGAAAGTACAGATCCATGACAAATCTGTTCTTCCTTACGCGTACTTACACC 1210

DB 4 CAGAAGAGAGCGTACGAGCCCATGGAATTTGGCTCTCCCTTCACTGATGACCCC 63

QY 1211 AATTAGACTCTGAGAAATGGCCATCAATAGATGGCTGAGACACCAAAAGTAA 1266

DB 64 AAACAGACACTGAGATGGCCATCAACCGATGGCTGGAGACGACGAGAAAGTAA 119

RESULT 7 AA912881/c 82 bp mRNA linear EST 26-AUG-1998

LOCUS AA912881 82 bp mRNA linear EST 26-AUG-1998

DEFINITION C127D06.81 Soares_NFL_T_GBC_ST Homo sapiens cDNA clone

IMAGE:1524659 3' similar to TR:023121 023121 F10G19.3 PROTEIN. ;

mRNA sequence.

ACCESSION AA912881

VERSION AA912881

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

NCI-CCGAP http://www.ncbi.nlm.nih.gov/ccgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgarbs-r@mail.nih.gov

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.lnl.gov) for further information.

Insert Length: 509 Std Error: 0.00

Seq primer: -40ml3 fwd. ET from Amersham

High quality sequence stop: 1.

Location/Qualifiers

1. .82

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:1524659"

/lab_host="DH10B"

/clone_1lb="Soares_NFL_T_GBC_ST"

/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with

a modified polylinker; Site 1: Not I; Site 2: Eco RI;

Equal amounts of plasmid DNA from three normalized

libraries (fetal lung Bhl19W, testis NHT, and B-cell

NCI CGAP GCBI) were mixed, and as circles were made in

vitro. Following HAP purification, this DNA was used as

tracer in a subtractive cDNA hybridization reaction. The driver

was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools were collected of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 728096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 6.2%; Score 78.8; DB 1; Length 82;

Best Local Similarity 97.6%; Pred. No. 3.6e-11;

Matches 80; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1044 ATTCCTGATGAAATTTATATGTCATATGAGAACTTATGAAAGATCGGTATCGCA 1103

DB 82 ATTCCTGATGAAATTTATATGTCATATGAGAACTTATGAAAGATCGGTATCGCA 23

QY 1104 TCAGATGGCTATTCATATGAAA 1125

DB 22 TCAGTGGCTATTCATATGAAA 1

RESULT 8 BB323974 191 bp mRNA linear EST 11-JUL-2000

LOCUS BB323974 191 bp mRNA linear EST 11-JUL-2000

DEFINITION BB323974 RIKEN full-length enriched, 4 days neonate male adipose

Mus musculus cDNA clone B430102G11 3', mRNA sequence.

ACCESSION BB323974

VERSION BB323974.1 GI:9032288

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 191)

TITLE
JOURNAL
COMMENT

RIKEN Mouse ESTs (Konno, H., et al.)
Unpublished (2000)

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Teurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagao, S., Sasaki, N., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
Thermoregulation and thermoregulation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh, M., Katsunai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y., and Hayashizaki, Y.
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Please visit our web site (http://genome.rtc.riken.go.jp) for further details.

FEATURES
source

Location/Qualifiers
1. 191
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="B430102G11"
/sex="male"
/tissue_type="adipose"
/dev_stage="4 days neonate"
/lab_host="DH10B"
/clone_id="RIKEN full-length enriched, 4 days neonate male adipose"
/note="Site 1: Sali; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5']
GAGGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTT 3', cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 229.0. Second strand cDNA was prepared with the primer adapter of sequence [5'] GAGGAGAGATTCGAGTTATTAATATCCCCCCCCCC 3'. cDNA was cleaved with XhoI and BamHI. Vector: a modified Bluescript KS(+) after bulk excision from Lambda FLX I."

ORIGIN
Query Match

5.6%; Score 71.4; DB 2; Length 191;

Best Local Similarity 74.4%; Pred. No. 6.3e-09; Matches 90; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 1151 CAAAAAGAACCCAGCAAGTCCATGCAATCTGTTCTTCCCTTACGGTACTTACACC 1210
2 CAAAGAGAGACCTTACAGAGCCCAAGTACAGATTGCTCTTCCCTTACGGTACTGACCC 61
Db 1211 AATAGACTCTGAAATGCGCATCATGATGCTGAGACACACCAAGTAAAGAT 1270
62 CACAGAGACCTGAAAGTGGCTTCAACGATGCTGAGACCCACGAGAGAGGAGCGT 121
QY 1271 T 1271
Db 122 T 122

RESULT 9
BI186455
LOCUS

BI186455 RIKEN full-length enriched, adult male spinal cord Mus
musculus cDNA clone A330033K06 3', mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

EST.
Mus musculus (house mouse)

REFERENCE
AUTHORS

1 (bases 1 to 177)
Konno, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M., Izawa, M., Kadoya, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shingawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomimaga, N., Toya, T., Tsunoda, Y., Watanabe, S., Yamamura, T., Yamanka, I., Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.
RIKEN Mouse ESTs (Konno, H., et al.)
Unpublished (2000)

TITLE
JOURNAL
COMMENT

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Teurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagao, S., Sasaki, N., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
Thermoregulation and thermoregulation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh, M., Katsunai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y., and Hayashizaki, Y.
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Please visit our web site (http://genome.rtc.riken.go.jp) for further details.

FEATURES
source

Location/Qualifiers
1. 177
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="A330033K06"

```
/sex="male"
/tissue_type="spinal cord"
/dev_stage="adult"
/lab_host="DH108"
/clone_lib="RIKEN full-length enriched, adult male spinal
cord"
/notes="Site 1: Sali; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5']
GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTTAA 3', cDNA was
prepared by using triphase thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 10.0 and subtraction to Rot = 459.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5'] GAGAGAGAGATTCGAGTTAAATTAATCCCCCCCCCCC
3']. cDNA was cleaved with XhoI and BamHI. Vector: a
modified Bluescript KS(+) after bulk excision from Lambda
FLC I."
```

ORIGIN

```
Query Match 3.9%; Score 49.2; DB 2; Length 177;
Best Local Similarity 73.3%; Pred. No. 0.014;
Matches 63; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
```

```
QY 1181 TCTTGTCTTCTTCACGGCTACTTACCAATAGACTCTGAATGGCATCAATAG 1240
Db 18 TTGCTTTTCTCTCTACTCTCCACCCCAACAGACACTGAAGTCCCATCAACG 77
QY 1241 ATGCTGAGAGACACCAAAAGTAA 1266
Db 78 ATGCTGAGAGACCAAGAACTGAA 103
```

RESULT 10

```
CR455364 102 bp mRNA linear EST 02-JUL-2004
LOCUS CR455364 Day 14 bovine embryos (bcal) Bos taurus cDNA clone
DEFINITION bcal0015a.h.02 5', mRNA sequence.
ACCESSION CR455364
VERSION CR455364.1 GI:49623753
KEYWORDS EST.
```

SOURCE

```
ORGANISM Bos taurus (cow)
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovinae; Bos.
```

```
REFERENCE 1 (bases 1 to 102)
AUTHORS Renard,J.P., Sreenan,J.P. and Hue,I.
TITLES Embryonic ESTs (bcal)
JOURNAL Unpublished (2004)
COMMENT Contact: Renard JP
Biologie du Developpement et Reproduction
INRA
```

Domaine de Vilvert 78350 Jouy en Josas, FRANCE
Email: renard@jouy.inra.fr

Funding for EST sequencing was provided by an INRA funding (AIP00183). Funding for the cDNA library construction was provided by an EEC contract (BO14-CT95-0190).

The library was constructed by I. Hue, BDR, INRA, 78350 JOUY-EN-JOSAS. The double stranded cDNA was unidirectionally cloned in a TA cloning vector. bcal is a SMART cDNA library. Insert size was between 450 and 3500 bp. Sequencing was done from the 5' end of the clone.

Sequence cleaned of vector, adaptor and repetitions. Contact us at esgenasupport@jouy.inra.fr to obtain the chromatogram of this sequence.

FEATURES
source Location/Qualifiers
1..102

ORIGIN

```
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/clone="bcal0015a.h.02"
/tissue_type="early elongating embryos (+/- 15 mm)"
/dev_stage="embryon"
/lab_host="DH5"
/clone_lib="Day 14 bovine embryos (bcal)"
/notes="Vector: PCR2.1 (Invitrogen); Clone distribution :
AGENAE Resource centre, Francois PUMI,
Francois.Pumi@jouy.inra.fr, INRA, CEA Radiobiologie et
Etude du genome (LREG), Domaine de Vilvert, 78352,
Jouy-en-Josas cedex, FRANCE, +33 (0) 1.34.65.28.02, +33
(0) 1.34.65.22.73"
```

```
Query Match 3.6%; Score 45.4; DB 7; Length 102;
Best Local Similarity 87.3%; Pred. No. 0.15;
Matches 62; Conservative 0; Mismatches 6; Indels 3; Gaps 1;
```

```
QY 1199 GGTACTTACACCAATAGAGCTGTGAATGGCATCAATAGTGGCTG---GAGACACA 1255
Db 1 GATCTTACACCAATAGAGCTGTGAATGGCATCAATAGTGGCTGCTAAGTACACACA 60
```

```
QY 1256 CCAAAAGTAA 1266
Db 61 TCAAAATATA 71
```

RESULT 11

```
AW032239 189 bp mRNA linear EST 18-MAY-2001
LOCUS AW032239
DEFINITION EST275693 tomato callus, TAMU Lycopersicon esculentum cDNA clone
CLEC35P22, mRNA sequence.
```

```
ACCESSION AW032239
VERSION AW032239.1 GI:5890995
KEYWORDS EST.
```

SOURCE Lycopersicon esculentum (tomato)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanales; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE 1 (bases 1 to 189)
Alcala,J., Vrebalov,D., White,R., Matern,A.L., Vision,T.,
Holt,I.B., Liang,F., Upson,J., Craven,M.B., Bowman,C.L., Ahn,S.,
Romling,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and
Giovannoni,J.

Generation of ESTs from tomato callus tissue
Contact: CUGI
Clemson University Genomics Institute

100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
5 prime sequence.

FEATURES
source Location/Qualifiers
1..189

/organism="Lycopersicon esculentum"

/mol_type="mRNA"

/cultivar="TA496"

/db_xref="taxon:4081"

/clone="CLEC35P22"

/tissue_type="callus"

/dev_stage="25-40 days old"

/lab_host="XLI-Blue MFP"

/clone_lib="tomato callus, TAMU"

/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Giovannoni laboratory; CLEC - Cotyledons
of seedlings 7-10 days post-germination were excised, cut
at both ends and placed on MS medium with no selection.
Mixed callus was harvested at 25 and 40 days and included
undifferentiated masses. Tomato Callus EST Library"

ORIGIN

Query Match 3.4%; Score 42.8; DB 2; Length 189;
Best Local Similarity 61.8%; Pred. No. 1;
Matches 68; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 1043 AATTCCTGATGAATTTATATGTCTCCAACTAGAGAACTTTGAAAGATCCGGTCATCC 1102
DB 74 AGTACCCCTTATTTATTTTCCCGCATTTCTTAAAGATGAAAGATCCGTCCAGCAT 133

QY 1103 ATCAGATGCTATTCATATGAAGAAGCATGGAATGATCAGCA 1152
DB 134 CTCACACTGGAGATGACATATGATCGGAAACATCGAAGATGATTTCA 183

RESULT 12
AG211342 184 bp DNA linear GSS 09-AUG-2003
LOCUS Oryza sativa (japonica cultivar-group) DNA, clone:NP4028.0.703.1A,
DEFINITION 3' flanking sequence of Tos17 insertion in rice strain NF4028,
genomic survey sequence.
ACCESSION AG211342 GI:32358532
VERSION AG211342.1 GI:32358532
KEYWORDS Oryza sativa (japonica cultivar-group)
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehharitoidae; Oryzaceae; Oryza.

REFERENCE 1 Miyao, A., Tanaka, K., Murata, K., Sawaki, H., Takeda, S., Abe, K.,
AUTHORS Shinzuka, Y., Onosato, K. and Hirochika, H.
TITLE Target Site Specificity of the Tos17 Retrotransposon Shows a
Preference for Insertion within Genes and against Insertion in
Retrotransposon-Rich Regions of the Genome
JOURNAL Plant Cell 15 (8), 1771-1780 (2003)
MEDLINE 22779046
PUBMED 12897251

REFERENCE 2 (bases 1 to 184)
AUTHORS Miyao, A., Sawaki, M. and Hirochika, H.
TITLE Direct Submission
JOURNAL Submitted (16-APR-2002) Akio Miyao, National Institute of
Agrobiological Sciences, Molecular Genetics; 2-1-2, Kannondai,
Tsukuba, Ibaraki 305-8602, Japan (E-mail:miyaoaffrc.go.jp,
URL:http://cos.nias.affrc.go.jp/, Tel:81-298-38-7020,
Fax:81-298-38-7020)

FEATURES
Location/Qualifiers
source
1..184
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/strain="NP4028"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="NP4028.0.703.1A"
/clone_1lb="PCR product directly amplified from rice
genomic DNA"
/note="The 3' end of retrotransposon Tos17 was found
immediately upstream of this sequence."

ORIGIN
misc_feature
1 /note="This insertion point is base 115457 in AC020666."

Query Match 3.1%; Score 39.8; DB 9; Length 184;
Best Local Similarity 59.1%; Pred. No. 7.3;
Matches 68; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 1066 CAATTAAGTGAAGCTTATGAAGATCCGATCGATCGATGCGTATTATCATATGAAA 1125
DB 5 CAATATTGAGAGAGTATGAGAGATCTCATTCGACGAGATGGTTCAGTACAGAG 64

QY 1126 AGAAGCAATGGAATTTGATCGCAAAAAGAAAGCTGCAAGTCCATGCAAAA 1180
DB 65 AGAGGCGCATCAAGAGATGCTCGGACGCGGCGACACACCTCCCGCATGACAAA 119

RESULT 13
A1622018/c 144 bp mRNA linear EST 22-APR-1999
LOCUS A1622018
DEFINITION 466031B09.x4 486 - leaf primordia cDNA library from Hake 1ab Zea
mays cDNA, mRNA sequence.
ACCESSION A1622018
VERSION A1622018
KEYWORDS A1622018.1 GI:4646943
SOURCE EST.
ORGANISM Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 144)
Malbot, V.
REFERENCE
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
AUTHORS
JOURNAL Unpublished (1999)
COMMENT Contact: Malbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: malbot@stanford.edu
Plate: 486031 row: E column: 09.

FEATURES
Location/Qualifiers
source
1..144
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="B73"
/db_xref="taxon:4577"
/issue_type="leaf primordia"
/dev_stage="P7-P11 leaf"
/lab_host="E.coli XL1-Blue MFR"
/clone_1lb="486 - leaf primordia cDNA library from Hake
1ab"
/note="Organ: shoot; Vector: Lambda zap; Hake 1ab cDNA
library."

ORIGIN
Query Match 3.1%; Score 39.2; DB 1; Length 144;
Best Local Similarity 58.6%; Pred. No. 9.9;
Matches 68; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 1041 GGAATTCCTGATGAATTTATATGCTCAATACAGAACTTATGAAGATCCGTCATC 1100
DB 143 GCACCTCGAGGCATTTTCATCTGCGCCATCTTCAGGAAGTATGTCATCCATAGCTC 84

QY 1101 GCATCAATGCGCTATTTATATGAAAAGCAATGGAATTTGATCAGCAAAA 1156
DB 83 GCGTCAGATGCTACACCTACGACGCGACGACCATTTAGCTGTGGCTGACGACGAA 28

RESULT 14
BE240075 175 bp mRNA linear EST 12-JUN-2000
LOCUS BE240075
DEFINITION BE2404124 MHRP- Medicago truncatula cDNA clone pMHRP-41H18, mRNA
sequence.
ACCESSION BE240075
VERSION BE240075
KEYWORDS BE240075.1 GI:9056258
SOURCE EST.
ORGANISM Medicago truncatula (barrel medic)
Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoidae; Trifoliaceae;
Medicago.
1 (bases 1 to 175)
Harrison, M. J., Town, C. D., Bowman, C. L., Craven, M. B., Hansen, T. S.,
Holt, I. E., Cho, J. and Fraser, C. M.
ESTs from phosphate-starved roots of Medicago truncatula

JOURNAL COMMENT

Unpublished (2000)
Contact: Maria J. Harrison
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73401, USA
Tel: 580-223-5810
Fax: 580-221-7380
Email: mjharrison@noble.org
The Samuel Noble Roberts Foundation: N265942e
TIGR sequence name: MTHAL45TK
More information is available at:
<http://chryslie.tamu.edu/medicago>
Seq primer: Skmod (CTA gaa cta gtc gat cc).
Location/Qualifiers

FEATURES

source

```
1..175
/organism="Medicago truncatula"
/mol_type="mRNA"
/cultivar="A17"
/db_xref="taxon:3880"
/clone="pMHRP-41H18"
/tissue_type="roots"
/dev_stage="phosphate-starved"
/lab_host="XtLR"
/clone_1lb="MHRP-"
/notes="Vector: pBluescript SK-; Site 1: EcoRI, Site 2:
XhoI; At the trifoliolate stage, M. truncatula plants were
transplanted to phosphate-free sand and grown for a
further 30 days. During this period, they were fertilized
twice weekly with 1/2 Hoagland solutions containing 200M
potassium phosphate. cDNA was prepared from polyA+
enriched RNA. The cDNA was directionally ligated into the
unizap XR vector from Stratagene and packaged using
Gigapack III Gold packaging extracts. Plasmids containing
cDNA inserts were excised from the recombinant lambda-Zap
phage using Ex-aseIst helper phage and propagated in
XtLR cells."
```

ORIGIN

Query Match

Best Local Similarity 2.7%; Score 34.2; DB 2; Length 175;
Matches 72; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

```
QY 1046 TCCTGATGAATTTATATGTCGAATTAAGCACTTATGAAGATCCGTCATCCGATC 1105
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 41 TCCTGATGAATTTATATGTCGAATTCCTCGAAGATCATGAGATCAGATGATGATGC 100
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1106 AGATGCTATTCATATATAAGAGCAATGGAATTTGATCAGCAAAAAGAACTTAC 1165
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 101 AGATGCTTTTACTTACAGACAGAGGCTTCCGAGATGCTTGAACAGTGCATGACGC 160
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1166 AAGTCCCATGACAAA 1180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 161 TTCACCGATGACAAA 175
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 15

CF370509

LOCUS CF370509 195 bp mRNA linear EST 27-AUG-2003
DEFINITION rgs05008.v1 Meloidogyne hapla female SMART pGEM Meloidogyne hapla
ACCESSION CF370509
VERSION CF370509.1 GI:34317755
KEYWORDS EST
SOURCE Meloidogyne hapla
ORGANISM Meloidogyne hapla
Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
Tylenchoidea; Heteroderidae; Meloidogyinae; Meloidogyne.

REFERENCE

AUTHORS

McCarter,J., Clifton,S., Chipelli,B., Pape,D., Martin,J.,
Wyle,T., Dence,M., Marra,M., Hillier,U., Kucaba,T., Theising,B.,
Bowers,Y., Gibbons,M., Ritzer,E., Bennett,J., Franklin,C.,
Teagarshevili,I., Ronko,I., Kennedy,S., Maguire,L., Beck,C.,
Underwood,K., Stepcoe,M., Allen,M., Person,B., Swaller,T.,

TITLE
JOURNAL COMMENT

Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McCam,R., Waterston,R. and Wilson,R.
The Washington Univ. Nematode EST Project, 1999
Unpublished (1999)
Contact: McCarter JP
The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
Cloned unidirectionally. Poly(A)⁺ RNA was concentrated and purified
using Dynabeads (Dyna) and mRNA eluted for first strand synthesis.
First strand cDNA was created using MMLV RT (Powerscript, Clontech)
and primed with oligo(dT) with XhoI site and 5' SMART 'anchor' added
using chimeric DNA-RNA oligo. 12 PCR cycles were done using first
strand and primers specific to SMART oligo and 3' end. Double
stranded cDNA was digested using XhoI/NotI, fractionated on
chroma-spin 400 columns (Clontech) and ligated to digested
pGEM-11zf(+) plasmid. Chemically competent DH10B cells were used as
host cells. Library materials provided by Dr. David Bird of North
Carolina State University. Library construction by Jeff Rousch. See
www.nematode.net for additional project information.
Putative full length read
The vector to vector length is 196
Seq primer: Sp6.

FEATURES

source

```
1..195
/organism="Meloidogyne hapla"
/mol_type="mRNA"
/db_xref="taxon:6305"
/tissue_type="whole organism"
/dev_stage="adult"
/lab_host="DH10B"
/clone_1lb="Meloidogyne hapla female SMART pGEM"
/notes="Vector: plasmid (ampicillin resistant); Site 1:  
XhoI; Site 2: NotI; Cloned unidirectionally. Poly(A)+ RNA  
was concentrated and purified using Dynabeads (Dyna) and  
mRNA eluted for first strand synthesis. First strand cDNA  
was created using MMLV RT (Powerscript, Clontech) and  
primed with oligo(dT) with XhoI site and 5' SMART 'anchor'  
added using chimeric DNA-RNA oligo. 12 PCR cycles were  
done using first strand and primers specific to SMART  
oligo and 3' end. Double stranded cDNA was digested using  
XhoI/NotI, fractionated on Chroma-spin 400 columns  
(Clontech) and ligated to digested pGEM-11zf(+) plasmid.  
Chemically competent DH10B cells were used as host cells.  
Library materials provided by Dr. David Bird of North  
Carolina State University. Library construction by Jeff  
Rousch. See www.nematode.net for additional project  
information."
```

ORIGIN

Query Match

Best Local Similarity 2.7%; Score 33.8; DB 7; Length 195;
Matches 62; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

```
QY 866 AGATTTAAAGATCTGTTGATTTTCAAGATGAATTAACATTTGATGAGAAAGAACTGTT 925
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 58 AATTAATAAAATTTTATTTATATGTTTCAATTTTAAAGAAATTTTAAAGAAATTTTAAAG 117
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 926 GAATCTTACAAAGAAAGTCTGCTGATGATTTGAATTTGAATCTCTA 974
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 118 AATTTTCAAAATTAATAATTTTCAAGATTTTGAATTTGACTTTTAA 166
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: February 5, 2005, 11:39:03
Job time : 4600 secs

;; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
;; FILE REFERENCE: 107196.132
;; CURRENT APPLICATION NUMBER: US/09/248, 796A
;; PRIOR FILING DATE: 1999-02-12
;; PRIOR APPLICATION NUMBER: US 60/074, 725
;; PRIOR FILING DATE: 1998-02-13
;; PRIOR APPLICATION NUMBER: US 60/096, 409
;; NUMBER OF SEQ ID NOS: 28208
;; SEQ ID NO 13675
;; LENGTH: 189
;; TYPE: DNA
;; ORGANISM: Candida albicans
US-09-248-796A-13675

Query Match 2.2%; Score 28.4; DB 4; Length 189;
Best Local Similarity 48.7%; Pred. No. 79;
Matches 77; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 931 TTACAAAAGAAAGCTGCTGATGATTTGAAATTTGATCTAGAGCTGCTAGTAAAG 990
DB 183 TTGAAAAAAATGATCTGAGAGGGGAGAAAGGGGAAATTAATTTGAGTTAAAAAAA 124
QY 991 TGCTGAGAAATTTGAAGAGCTCAGACCAAGTTAAATCCCTTTCTTCAGGAATTCCTG 1050
DB 123 TAAAGGTAAAGGTGAAAGGTGAAAGGGTGGGTTTGGTTTGTGTAATCCATTTCTT 64
QY 1051 ATGAAATTTATATGCTCAATTAAGTAACTAGAGAACTTATGAAA 1088
DB 63 ATCTCTATATTAATTCATTAATTAATTAATTAATTA 26

RESULT 3
US-09-513-999C-31162
;; Sequence 31162, Application US/09513999C
;; Patent No. 6783961
;; GENERAL INFORMATION:
;; APPLICANT: Dumas Milne Edwards, J.B.
;; APPLICANT: Duclert, A.
;; APPLICANT: Giordano, J.Y.
;; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
;; Patent No. 6783961
;; FILE REFERENCE: 59.US2.REG
;; CURRENT APPLICATION NUMBER: US/09/513, 999C
;; CURRENT FILING DATE: 2000-02-24
;; PRIOR APPLICATION NUMBER: US 60/122, 487
;; PRIOR FILING DATE: 1999-02-26
;; NUMBER OF SEQ ID NOS: 36681
;; SOFTWARE: Patent.pm
;; SEQ ID NO 31162
;; LENGTH: 194
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-513-999C-31162

Query Match 2.2%; Score 28.2; DB 4; Length 194;
Best Local Similarity 49.7%; Pred. No. 93;
Matches 72; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 860 TGCAAGATTTTAAAGATCTTGTGTTATTTTCAAGATGATTAACCTTGATGAAAAGA 919
DB 1 TGCAATCCCATTTAAGAAACCTGTAACTTGAAGATAGAGACTTAATCATCAAAACAGA 60
QY 920 ACTGTTGAATCTTCAAAAAGAAAGTCTGGCTGATGATTTGAAAATTTGAATCTTAGAGACT 979
DB 61 ACAGATGGAACAGATTTAAAGAAACAGAGACAACTTACAAATTAAGAAATCTCAAGAGA 120
QY 980 GCGTAGTAAAGTCTGAGAAATTT 1004
DB 121 CGCAATAAAGATTAAGAAATCTT 145

RESULT 4

US-09-248-796A-8503/C
;; Sequence 8503, Application US/09248796A
;; Patent No. 6747137
;; GENERAL INFORMATION:
;; APPLICANT: Keith Weinstock et al
;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
;; FILE REFERENCE: 107196.132
;; CURRENT APPLICATION NUMBER: US/09/248, 796A
;; CURRENT FILING DATE: 1999-02-12
;; PRIOR APPLICATION NUMBER: US 60/074, 725
;; PRIOR FILING DATE: 1998-02-13
;; PRIOR APPLICATION NUMBER: US 60/096, 409
;; NUMBER OF SEQ ID NOS: 28208
;; SEQ ID NO 8503
;; LENGTH: 183
;; TYPE: DNA
;; ORGANISM: Candida albicans
;; FEATURE:
;; NAME/KEY: unsure
;; LOCATION: (11)
;; OTHER INFORMATION: Identity of nucleotide sequences at the above locations are unknown
US-09-248-796A-8503

Query Match 2.2%; Score 27.6; DB 4; Length 183;
Best Local Similarity 52.2%; Pred. No. 1,4e+02;
Matches 60; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 515 TAGGTGTAAGAGGCTCTGCGGCGCATGATTTCTCCTTAATGAAAGCTCTTGT 574
DB 117 TATTTCAATGCTGCTATTTTGTCTGCATTAATTAATAGTGGCTAAATGCTAGTGTGTC 58
QY 575 CACTGGCTCCTCATGTGTGATTTTACAGTGTGATGATTAAGTAAATGAGGTCTG 629
DB 57 TACATCAAGCCATTGATGCTTACACAGAGGCGCATGGAACAAAGTTTCTG 3

RESULT 5
US-08-875-972-27/C
;; Sequence 27, Application US/08875972
;; Patent No. 598564
;; GENERAL INFORMATION:
;; APPLICANT: Huntington Potter and Jinhue Li
;; TITLE OF INVENTION: ASSAY FOR IDENTIFYING GENES CAUSING
;; TITLE OF INVENTION: CHROMOSOME NON-DISJUNCTION
;; NUMBER OF SEQUENCES: 29
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
;; STREET: Two Millitia Drive
;; CITY: Lexington
;; STATE: Massachusetts
;; COUNTRY: USA
;; ZIP: 02173-4799
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Releasee #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/875, 972
;; FILING DATE: 08-AUG-97
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 60/002, 448
;; FILING DATE: 16-AUG-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Granahan Esq., Patricia
;; REGISTRATION NUMBER: 32,227
;; REFERENCE/DOCKET NUMBER: HU95-03PA
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (781) 861-6240
;; TELEFAX: (781) 861-9540

INFORMATION FOR SEQ ID NO: 27;
SEQUENCE CHARACTERISTICS:
LENGTH: 200 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-875-972-27

Query Match 2.2%; Score 27.6; DB 2; Length 200;
Best Local Similarity 49.3%; Pred. No. 1.5e+02;
Matches 72; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 1056 TTTATATGTCGAATAGAGCACTTATGAAAGATCCGTCATCGCATGAGTGGCAT 1115
DB 181 TTTATGGGCTCAAAATTTCTGTGACAAATTTTGGTCAAGTTGTTCCATTAAAAAGTAC 122
QY 1116 TCATATGAAAAAGAAAGCAATGGAATTTGATCAGCAAAAGAAACGTACAGTCCCATG 1175
DB 121 TGATTTTAAACCTAATACTTAAGCTGACACGCAAGAAAAAGAAACCAAGTGTGTC 62
QY 1176 ACAATCTGTTCTTCTTCCTCAGCGGT 1201
DB 61 ACAAAACATCTCTGCTTCTGAGGT 36

RESULT 6
US-09-328-750A-5
Sequence 5, Application US/09328750A
Patent No. 6410722
GENERAL INFORMATION:
APPLICANT: McGill University et al.
TITLE OF INVENTION: HUMAN AND MAMMALIAN DNA
REPLICATION ORIGIN CONSENSUS SEQUENCES
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: SMABEY OGILVY RENAULT
STREET: 1981 McGill College Avenue - Suite 1600
CITY: Montreal
STATE: QC
COUNTRY: Canada
ZIP: H3A 2Y3
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/328,750A
FILING DATE: 09-Jun-1999
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/033,374
FILING DATE: 16-DEC-1996
APPLICATION NUMBER: 60/047,322
FILING DATE: 21-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: CUC, France
REGISTRATION NUMBER: 4166
REFERENCE/DOCKET NUMBER: 1770-162PCT FC/1d
TELECOMMUNICATION INFORMATION:
TELEPHONE: 514 845-7126
TELEFAX: 514 288-8389
INFORMATION FOR SEQ ID NO: 5;
SEQUENCE CHARACTERISTICS:
LENGTH: 91 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
SEQUENCE DESCRIPTION: SEQ ID NO: 5;
US-09-328-750A-5

Query Match 2.1%; Score 27; DB 3; Length 91;
Best Local Similarity 15.7%; Pred. No. 1.4e+02;
Matches 13; Conservative 49; Mismatches 21; Indels 0; Gaps 0;

QY 871 TAAAGATCTGTTGTAATTTCAAGATGATTAATCATTTGAGAAAGAACTGTAATC 930
DB 8 KRAMRWKXDAVWVGARWKKVWHRASSACMDKAKATWKGWTAARRYWKGRKMWWTW 67
QY 931 TTACAAAAGAAAGTCTGCTGAT 953
DB 68 KAMSDATAYKXWMDAKKMWKRT 90

RESULT 7
US-08-369-829A-13
Sequence 13, Application US/08369829A
Patent No. 5861377
GENERAL INFORMATION:
APPLICANT: Hans Fritz
APPLICANT: Christian Sommerhoff
APPLICANT: Jutta Heim
TITLE OF INVENTION: No. 5861377el Inhibitor
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5861377artis Corporation
STREET: 59 Route 10
CITY: East Hanover
STATE: New Jersey
COUNTRY: US
ZIP: 07936
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25 (BPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/369,829A
FILING DATE: 6-JAN-95
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EPO 94810006.0
FILING DATE: 7-JAN-94
ATTORNEY/AGENT INFORMATION:
NAME: No. 5861377ak, Henry P.
REGISTRATION NUMBER: 33200
REFERENCE/DOCKET NUMBER: 4-19942/A/DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 277-5110
TELEFAX: (908) 277-4306
INFORMATION FOR SEQ ID NO: 13;
SEQUENCE CHARACTERISTICS:
LENGTH: 90 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..90
OTHER INFORMATION: /function= "synthetic oligo for PCR"
US-08-369-829A-13
Query Match 2.1%; Score 26.8; DB 2; Length 90;
Best Local Similarity 61.4%; Pred. No. 1.6e+02;
Matches 43; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
QY 1172 CATGCAAAATCTGTTCTTCAGCGGTAAGTACCAAAATGAGACTGAAATGCG 1231
DB 6 CACAACCGTTTCTGCTCTTCAACCGTAATTACACGAATACGACAGTAACTTCTGT 65
QY 1232 CATCATAGA 1241
DB 66 TACAAACACA 75

```

RESULT 8
US-09-513-999C-35764/C
: Sequence 35764, Application US/0951399C
: Patent No. 6783961
: GENERAL INFORMATION:
: APPLICANT: Dumas Milne Edwards, J.B.
: APPLICANT: Duclert, A.
: APPLICANT: Giordano, J.Y.
: TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
: Patent No. 6783961
: FILE REFERENCE: 59. US2.REG
: CURRENT APPLICATION NUMBER: US/09/513,999C
: CURRENT FILING DATE: 2000-02-24
: PRIOR APPLICATION NUMBER: US 60/122,487
: PRIOR FILING DATE: 1999-02-26
: NUMBER OF SEQ ID NOS: 36681
: SOFTWARE: Patent.pm
: SEQ ID NO 35764
: LENGTH: 200
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-513-999C-35764

```

Query Match	2.1%	Score 26.8;	DB 4;	Length 200;
Best Local Similarity	59.0%;	Pred. No. 2.8e+02;		
Matches 46;	Conservative 0;	Mismatches 32;	Indels 0;	Gaps 0

Qy	858	TGTGACAA	GATT	TA	AAAG	ATCT	TGTG	TA	TTTTC	CAAG	AT	CA	TTG	ATG	GA	917	
Db	137	TAT	CTAA	TAG	TTT	CTG	CA	TTT	AT	TTG	CT	T	TA	AA	AT	AT	78

Qy	918	GAACGTGTGAATCTTACA	935
Db	77	GAGTCCTCGACGTTTCA	60

RESULT 5

US-08-739-103A-12
; Sequence 12, Application US/08739103A
; Patent No. 6265154

GENERAL INFORMATION:
 APPLICANT: P. Kroeger
 APPLICANT: K. Abiravaya
 APPLICANT: J. Gorzowski
 APPLICANT: R. Hoehnle
 APPLICANT: J. Moore
 TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FOR DETECTING

```

? NUMBER OF SEQUENCES: 41
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Abbott Laboratories
? STREET: 100 Abbott Park Road
? CITY: Abbott Park
? STATE: Illinois
? COUNTRY: USA
?

```

```

?      ZIP: 60064-3500
?
?      COMPUTER READABLE FORM:
?
?      MEDIUM TYPE: Floppy disk
?
?      COMPUTER: Macintosh
?
?      OPERATING SYSTEM: System 7.0.1
?
?      SOFTWARE: Microsoft Word 5.1a
?
?      CURRENT APPLICATION DATA:
?

```

APPLICATION NUMBER: US/08/739,103A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Paul D. Yaeger
REGISTRATION NUMBER: 37,477
REFERENCE/DOCKET NUMBER: 5990.US.O1

TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/937-2341
TELEFAX: 847/938-2623

```

? TELEX:
? INFORMATION FOR SEQ ID NO: 12:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 139 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: genomic DNA (HPV type 33)
? IS-08-739-103A-12

```

Query Match	2.1%	Score 26.6	DB 3	Length 139
Best Local Similarity	53.3%	Pred. No. 2.5e+02		
Matches 55; Conservative	0	Mismatches 45	Indels 0	Gaps 0

843 GTCACACATGGCTTTGTGCACAAAGTTTAAAAGATCTTGTGGTATTTTCAAGTGAAT 902

903 AACATTGAGAAAGAACTGTTGAATCTTACAAAAGAAAGCTG 947
 904
 905
 906
 907
 908
 909
 910
 911
 912
 913
 914
 915
 916
 917
 918
 919
 920
 921
 922
 923
 924
 925
 926
 927
 928
 929
 930
 931
 932
 933
 934
 935
 936
 937
 938
 939
 940
 941
 942
 943
 944
 945
 946
 947
 948
 949
 950
 951
 952
 953
 954
 955
 956
 957
 958
 959
 960
 961
 962
 963
 964
 965
 966
 967
 968
 969
 970
 971
 972
 973
 974
 975
 976
 977
 978
 979
 980
 981
 982
 983
 984
 985
 986
 987
 988
 989
 990
 991
 992
 993
 994
 995
 996
 997
 998
 999
 1000

RESULT 10
US-09-248-796A-10429

Sequence 10429, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstock et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.132

1 CURRENT APPLICATION NUMBER: US/09/248,796A
 2
 3 CURRENT FILING DATE: 1999-02-12
 4
 5 PRIOR APPLICATION NUMBER: US 60/074,725
 6
 7 PRIOR FILING DATE: 1998-02-13
 8
 9 PRIOR APPLICATION NUMBER: US 60/096,409
 10
 11 PRIOR FILING DATE: 1998-08-13
 12
 13 NUMBER OF SEQ. ID NOS.: 2620A

```

1 SEQ ID NO 10429
2
3 LENGTH: 186
4
5 TYPE: DNA
6
7 ORGANISM: Candida albicans
8
9 US-09-248-796A-10429
10
11 Query Match 2.1% Score 26.4 DB 4 Length 186

```

867	ATGCGCCACCAAGCAATTTTTCCTTCCTCCTCATTTTTCAGAGAAATAACAATTGGACAGCAA	916
Match	48; Conservative	0; Mismatches 36; Indels 0; Gaps 0;
Best local similarity	57.1%; E=0.0	No. conserved

Db 72 TTGTACAAAGATTAAGAAATTCCTTGATGATTAATTAGTAAAGTATGCTAAT 131

Db
132 TGAATTTATTTTCACAAAAGA 155

RESULT 11
US-09-248-796A-9338/c
: Sequence 9338, Application US/09248796A
: Patent No. 6747137
: GENERAL INFORMATION:
: APPLICANT: Keith Weinstein et al

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
;

NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 9338
; LENGTH: 195
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-9338

Query Match 2.1%; Score 26.4; DB 4; Length 195;
Best Local Similarity 54.0%; Pred. No. 3.7e+02;
Matches 54; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 986 TAAAGTCCTGAGGAAATGGAAGCTCAGACCAAGTTAAATCCCTTCTTCAGAAAT 1045
195 TCAAGTGTGAGCAATTCGCGCAATTAACATCCAGAACCAATTTTCTAGTTAAATC 136

QY 1046 TCCGATGATTTATATGTCCTCAATACTAGAACTATG 1085
135 TTCTGTGAATTCATATTGTGATATGTCGAGAGATTGG 96

RESULT 12

US-09-248-796A-14068
; Sequence 14068, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248.796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 14068
; LENGTH: 198
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-14068

Query Match 2.1%; Score 26.4; DB 4; Length 198;
Best Local Similarity 59.2%; Pred. No. 3.7e+02;
Matches 45; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 900 AATACATGATGAGAAAGAACTGTGATCTTACAAAGAACTGCTGATGATTTG 959
22 AATATATGATGAGGAAACGCTAGTGAGATGAACAGAAATTTGATGAAGATGATTG 81

QY 960 AAAATGATCTCTAG 975
82 GATGATGATTTGAAG 97

RESULT 13

US-08-535-057A-11
; Sequence 11, Application US/08535057A
; Patent No. 6140104
; GENERAL INFORMATION:
; APPLICANT: LERECUS, DIDIER
; APPLICANT: AGAISSE, HERVE
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES FOR THE CONTROL OF
; TITLE OF INVENTION: THE EXPRESSION OF DNA SEQUENCES IN A CELL HOST
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/535,057A
FILING DATE: 20-DEC-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR94/00525
FILING DATE: 05-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 93/05387
FILING DATE: 05-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 660-103-0X PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 144 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..144
OTHER INFORMATION: /note= "CORRESPONDS TO NUCLEOTIDES
US-08-535-057A-11
OTHER INFORMATION: 1413 TO 1556 OF SEQ ID NO:1"

Query Match 2.1%; Score 26.2; DB 3; Length 144;
Best Local Similarity 54.7%; Pred. No. 3.5e+02;
Matches 52; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 1047 CCGATGATTTATATGTCCTCAATACTAGAGAACTTATGAAGAATCCGTCATGCATCA 1106
33 CATTAATAATATATATGACCGCTATATGATTAATAAATCATTTATCAGTTTG 92

QY 1107 GATGCTATTCATATGAAAGAAAGCAATGAGAA 1141
Db 93 AAATATATGATATATGATTAAGAAAGGAGAGAGAA 127

RESULT 14

US-09-503-252-11
; Sequence 11, Application US/09503252
; Patent No. 6553366
; GENERAL INFORMATION:
; APPLICANT: LERECUS, DIDIER
; APPLICANT: AGAISSE, HERVE
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES FOR THE CONTROL OF
; TITLE OF INVENTION: THE EXPRESSION OF DNA SEQUENCES IN A CELL HOST
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/503,252
 FILING DATE: 14-Feb-2000
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/535,057
 FILING DATE: <Unknown>
 APPLICATION NUMBER: FR 93/05387
 FILING DATE: 05-MAY-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: OBLON, NORMAN F.
 REGISTRATION NUMBER: 24,618
 REFERENCE/DOCKET NUMBER: 660-103-0X PCT
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 703-413-3000
 TELEFAX: 703-413-2220
 INFORMATION FOR SEQ ID NO: 11:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 144 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 FEATURE:
 NAME/KEY: misc.feature
 LOCATION: 1..144
 OTHER INFORMATION: /note="CORRESPONDS TO NUCLEOTIDES
 1413 TO 1556 OF SEQ ID NO:1"
 US-09-503-252-11
 SEQUENCE DESCRIPTION: SEQ ID NO: 11:

Query Match 2.1%; Score 26.2; DB 4; Length 144;
 Best Local Similarity 54.7%; Pred. No. 3.5e+02;
 Matches 52; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 1047 CCTGATGAATTTATGTCCTCAATTAAGAGAACTTATGAAGATCGGTCAGCATCA 1106
 DB 33 CATTAAAAACATATATTGACCGCTTAATGATTTATGAAAAATCAATTTATCAGTTTG 92
 QY 1107 GATGCTATTCAATGAGAAAGAACATGAGAAA 1141
 DB 93 AAATATTATGATATGATTAAGAAAGAGAGAGAA 127

RESULT 15
 US-09-134-001C-1737/c
 ; Sequence 1737, Application US/09134001C
 ; Patent No. 6380370
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
 ; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: GTC-007
 ; CURRENT APPLICATION NUMBER: US/09/134,001C
 ; CURRENT FILING DATE: 1998-08-13
 ; PRIOR APPLICATION NUMBER: US 60/064,964
 ; PRIOR FILING DATE: 1997-11-08
 ; PRIOR APPLICATION NUMBER: US 60/055,779
 ; PRIOR FILING DATE: 1997-08-14
 ; NUMBER OF SEQ ID NOS: 5674
 ; SEQ ID NO 1737
 ; LENGTH: 186
 ; TYPE: DNA
 ; ORGANISM: Staphylococcus epidermidis
 US-09-134-001C-1737

Query Match 2.1%; Score 26.2; DB 3; Length 186;
 Best Local Similarity 58.2%; Pred. No. 4.2e+02;
 Matches 46; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 895 AGATGATAATCATTTGATGAAAGAACTGTTGAATCTTACAAAGAAAGTGGCTGATG 954
 DB 93 AAATGAAAGACAAAGATTGATGAGTAACGAATGTTTAAAGTAGTTGAGATTACT 34

QY 955 ATTGAAAAATGATCTCT 973
 DB 33 ATAGAAAAAGCGTACCTCT 15
 Search completed: February 5, 2005, 11:43:11
 Job time: 247 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 5, 2005, 04:21:09 / Search time 5661 Seconds

(without alignments)
10887.658 Million cell updates/sec

Title: US-10-077-111-12

Perfect score: 1272

Sequence: 1 gaattcgagcttcacccgcg.....acacacaaagtaagaattc 1272

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: *
1: gb ba: *
2: gb hlg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vt: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1250.2	75.8	1564	9	BC029520 Homo sapi
2	964.2	75.8	1564	6	AR339277 Sequence
3	912.6	71.7	1254	6	CQ729621 Sequence
4	893.4	70.2	1919	9	AK129983 Homo sapi
5	890.2	70.0	1811	6	AX748164 Sequence
6	890.2	70.0	1811	9	AK093494 Homo sapi
7	878.8	69.1	1996	6	CQ491312 Sequence
8	878.8	69.1	1996	6	CQ497179 Sequence
9	478.6	37.6	1291	6	AX077675 Sequence
10	474.6	37.3	2130	10	BC061948 Mus muscu
11	436	34.3	1547	10	BC050792 Mus muscu
12	430.2	33.8	147021	9	AJ719461 Gallus ga
13	430	33.8	147021	9	AC006501 Homo sapi
14	430	33.8	177089	9	AC008277 Homo sapi
15	335	26.3	446	6	CQ473162 Sequence
16	331.6	26.1	366	6	CQ096643 Sequence
17	331.6	26.1	366	6	CQ135438 Sequence
18	331.6	26.1	366	6	CQ218732 Sequence
19	331.6	26.1	366	6	CQ294537 Sequence

C 20	331.6	26.1	366	6	CQ331208
C 21	329.8	25.9	409	6	CQ482331
C 22	328.8	25.8	441	6	CQ503452
C 23	328.8	25.8	441	6	CQ512287
C 24	297	23.3	297	6	CQ109842
C 25	297	23.3	297	6	CQ148540
C 26	297	23.3	297	6	CQ231810
C 27	297	23.3	297	6	CQ307124
C 28	297	23.3	297	6	CQ344013
C 29	276.8	21.8	245795	2	AC096419
C 30	276.8	21.8	315453	2	AC122070
C 31	275.4	21.7	196403	2	AC102767
C 32	275.4	21.7	209437	2	AC111119
C 33	275.4	21.7	217352	10	AL929242
C 34	179.6	14.1	466	6	CQ097537
C 35	179.6	14.1	466	6	CQ136375
C 36	179.6	14.1	466	6	CQ219679
C 37	179.6	14.1	466	6	CQ258151
C 38	179.6	14.1	466	6	CQ295485
C 39	179.6	14.1	466	6	CQ332079
C 40	179.6	14.1	153865	9	AC009307
C 41	176.8	13.9	180	6	CQ110706
C 42	176.8	13.9	180	6	CQ149443
C 43	176.8	13.9	180	6	CQ232723
C 44	176.8	13.9	180	6	CQ270754
C 45	176.8	13.9	180	6	CQ308036

ALIGNMENTS

RESULT 1	BC029520	1564 bp	mRNA	linear	PRI 20-JUL-2004
LOCUS	Homo sapiens WD repeat and SAM domain containing 1, mRNA (cDNA				
DEFINITION	clone MGC:33855 IMAGE:5301559), complete cds.				
ACCESSION	BC029520				
VERSION	BC029520.1	GI:20810486			
KEYWORDS	MGC.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 1564)				
AUTHORS	Strausberg, R.L., Peingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Helel, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Schaefer, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Adamson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, D., Myers, R.M., Butlerfield, Y.S., Krzywicki, M.I., Skalski, U., Smailus, D.E., Scherch, A., Schein, J.B., Jones, S.J. and Marra, M.A.				
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)				
PUBMED	12477932				
REFERENCES	2 (bases 1 to 1564)				
AUTHORS	Strausberg, R.				
TITLE	Direct Submission				
JOURNAL	Submitted (01-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550, USA				

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-ehgc.stanford.edu>
Contact: (Dickson, Mark) mcdpax11.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAC Plate: 48 Row: 0 Column: 11
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 22749102.
Location/Qualifiers

FEATURES

source

```
1..1564
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:33855 IMAGE:5301559"
/cisue_type="Brain, hypothalamus"
/clone_lib="NIH MGC_96"
/lab_host="DH10B"
/note="Vector: pBluescript"
1..1564
/clone="WDSAM1"
/note="synonym: FLJ36175"
/db_xref="LocusID:151525"
146..1300
/clone="WDSAM1"
/codon_start=1
/product="WDSAM1 protein"
/protein_id="AAH29520.1"
/db_xref="GI:20810487"
/db_xref="LocusID:151525"
/translation="MKYLITHTLADHDGDDVNCASFSLATSLDTRIRYSRDTPE  
LHSPLEFHTYAVHCCFSPSGHILASCSVDGTTVLMNTENGMVLVMEQPSGSPRV  
CQSPSPSTCLASGADGTIVLWNAQSKYLRCGSVXGSLAACAFAFNGSFFVYGS  
GDLTPVDDKRCLESHKADLGIITCCDFSPQVDSQGLAFAFNGSFFVYGS  
VSEPTDLARTHEHOLKOPTEDMESEEDVSTLCAQDLKDLVIGFRMNNIDKELNLT  
ESLADLKTESLGRSKVLRIKRIELRTKYLSSGIPEDFICITTEMLKEDPVIAADG  
YSYKRAMEMWISKRTSPMTLVLPNAVLTFRNTLKAIAINMLTHQK"
```

CDS

gene

ORIGIN

Query Match 98.3%; Score 1250.2; DB 9; Length 1564;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1252; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
QY 12 TCACCTGCGGCGACGTCACCGCCCGTGGGACCTTGAAGCGGATCCCGGCG 71
DB 47 TCACCTGCGGCGACGTCACCGCCCGTGGGACCTTGAAGCGGATCCCGGCG 106
QY 72 CCCCCGCTCTGACGCGTGTCTTCTCAATAAAGCACTGGTGAACCTGATTCACACA 131
DB 107 CCCCCGCTCTGACGCGTGTCTTCTCAATAAAGCACTGGTGAACCTGATTCACACA 166
QY 132 TTAGTGATCATGTGACGATGTCACACTGCTGCGCTTCTCTTTTCCCTTGGCTACT 191
DB 167 TTAGTGATCATGTGACGATGTCACACTGCTGCGCTTCTCTTTTCCCTTGGCTACT 226
QY 192 TGTCTCTTGGACAAACAATTGCGCTGACTCGTTACGTAAGCTGACCAT 251
DB 227 TGTCTCTTGGACAAACAATTGCGCTGACTCGTTACGTAAGCTGACCAT 286
QY 252 TCTCCATTGAAGTTTCACTAATGCTGTGACACTGCTGCTGTTCTCCCTTCAAGACAT 311
DB 287 TCTCCATTGAAGTTTCACTAATGCTGTGACACTGCTGCTGTTCTCCCTTCAAGACAT 346
```

```
QY 312 ATTTTGGCATCGTGTTCACACAGATGCTACACTGTCTTAATGAAATCTGAAAATGACAG 371
DB 347 ATTTTGGCATCGTGTTCACACAGATGCTACACTGTCTTAATGAAATCTGAAAATGACAG 406
QY 372 ATGCTGGCATGTATGGAACAAGCTTAATGGCAGCCCTGTAGAGGTTTGGCAATTTCCCA 431
DB 407 ATGCTGGCATGTATGGAACAAGCTTAATGGCAGCCCTGTAGAGGTTTGGCAATTTCCCA 466
QY 432 GACTCCACGTTTGGCATCAGGGGACGCTGATGGAATCTGTGGTTTGTGAAATGACAG 491
DB 467 GACTCCACGTTTGGCATCAGGGGACGCTGATGGAATCTGTGGTTTGTGAAATGACAG 526
QY 492 TCATCAAAATTATATAGATGTGTATGTATTAAGATGAGTCTCTTGGCGGATGTGCATT 551
DB 527 TCATCAAAATTATATAGATGTGTATGTATTAAGATGAGTCTCTTGGCGGATGTGCATT 586
QY 552 TCTCCTTAATGGAAGCTTTCTTTGTCACTGCTCTCATGTGTGATTTAACAGTGTGGAT 611
DB 587 TCTCCTTAATGGAAGCTTTCTTTGTCACTGCTCTCATGTGTGATTTAACAGTGTGGAT 646
QY 612 GATTAATGAGGTTGTGATAGTATGTAAGTATGTAAGTATGTAAGTATGTAAGTATGTAAG 671
DB 647 GATTAATGAGGTTGTGATAGTATGTAAGTATGTAAGTATGTAAGTATGTAAGTATGTAAG 706
QY 672 TTTTCTTCAAGCAGCTTTCTGATGAGAACAAAGTCTTCAATTTTTCGACTGCGATCA 731
DB 707 TTTTCTTCAAGCAGCTTTCTGATGAGAACAAAGTCTTCAATTTTTCGACTGCGATCA 766
QY 732 TGTGTGATGAGATTTGCCAAGTCAAAATTTGGATTTCTTTTACCATATCTTAGCAAG 791
DB 767 TGTGTGATGAGATTTGCCAAGTCAAAATTTGGATTTCTTTTACCATATCTTAGCAAG 826
QY 792 CGCACAGAAATCAGCTGTAAGCAATTTACCAAGATTTGTCAGAGAGGTCGCTCAACA 851
DB 827 CGCACAGAAATCAGCTGTAAGCAATTTACCAAGATTTGTCAGAGAGGTCGCTCAACA 886
QY 852 TGGCTTTGTGACCAAGATTTAAAGATCTTGTGTGATTTTTCAGATGATTAACATTTGAT 911
DB 887 TGGCTTTGTGACCAAGATTTAAAGATCTTGTGTGATTTTTCAGATGATTAACATTTGAT 946
QY 912 GGAAGAAAGCTGTTGAATCTTACAAAAGAAAGTCTGCGTGAATTTGAAAATTGAATCT 971
DB 947 GGAAGAAAGCTGTTGAATCTTACAAAAGAAAGTCTGCGTGAATTTGAAAATTGAATCT 1006
QY 972 CTAGACGCTGTTGTAAGTGTGAGAAATTTGAAAGCTCAGACCAAGGTTAAATCC 1031
DB 1007 CTAGACGCTGTTGTAAGTGTGAGAAATTTGAAAGCTCAGACCAAGGTTAAATCC 1066
QY 1032 CTTTCTTCAAGAAATTCCTGATGAAATTTATGTCCAAATACAGAACTTAAGAAAGAT 1091
DB 1067 CTTTCTTCAAGAAATTCCTGATGAAATTTATGTCCAAATACAGAACTTAAGAAAGAT 1126
QY 1092 CCGGTATGTCATCAGATGCTATTTATATGAAAAGAAAGCAATGGAATTTGATCAAC 1151
DB 1127 CCGGTATGTCATCAGATGCTATTTATATGAAAAGAAAGCAATGGAATTTGATCAAC 1186
QY 1152 AAAAAAGAAAGTACAGATGCTCCATGACCAATCTTTTCTTCCAGCGGTAATTAACCA 1211
DB 1187 AAAAAAGAAAGTACAGATGCTCCATGACCAATCTTTTCTTCCAGCGGTAATTAACCA 1246
QY 1212 AATAGACTCTGAAAATGSCCATCAATAGATGGCTGAGACACACCAAAAGTAA 1266
DB 1247 AATAGACTCTGAAAATGSCCATCAATAGATGGCTGAGACACACCAAAAGTAA 1301
```

RESULT 2
AR339277 1844 bp DNA linear PAT 17-AUG-2003
LOCUS AR339277
DEFINITION Sequence 768 from patent US 6569662.
ACCESSION AR339277
VERSION AR339277.1 GI:33726134
KEYWORDS
SOURCE Unknown.

ORGANISM

Unknown:
Unclassified.

1 (bases 1 to 1844)

AUTHORS Tang, Y.T., Zhou, P. and Drmanac, R.T.

TITLE Nucleic acids and polypeptides

JOURNAL Patent: US 6569662-A 2003-05-27-MAY-2003;

FEATURES

SOURCE 1.1844
/organism="genomic DNA"

ORIGIN

Query Match 75.8%; Score 964.2; DB 6; Length 1844;

Best Local Similarity 81.8%; Pred. No. 2.6e-250;

Matches 1252; Conservative 0; Mismatches 3; Indels 276; Gaps 1;

12 TCACCTGCGGCGACGTCGACCGCCCGTGGGCACTTGAAGCGCATCCCGGCG 71
 65 TCACCTGCGGCGACGTCGACCGCCCGTGGGCACTTGAAGCGCATCCCGGCG 124
 72 CCCCCGCTCCGACGCGCTTTCTTCAAAATAAGAAATGTAAGTGAATGATTCACACA 131
 125 CCCCCGCTCCGACGCGCTTTCTTCAAAATAAGAAATGTAAGTGAATGATTCACACA 184
 132 TTACCTGATGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT 191
 185 TTACCTGATGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT 244
 192 TGCTCTTGGACAAACAAATTCGCTGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT 251
 245 TGCTCTTGGACAAACAAATTCGCTGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT 304
 252 TCTCCATGTAAGTTCATACCTATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT 311
 305 TCTCCATGTAAGTTCATACCTATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT 364
 312 ATTTGGCATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT 371
 365 ATTTGGCATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT 424
 372 ATGCTGCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT 431
 425 ATGCTGCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT 484
 432 GACTCCAGTGTGGCATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT 491
 485 GACTCCAGTGTGGCATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT 544
 492 TCATACAAATTAATGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT 551
 545 TCATACAAATTAATGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT 604
 552 TCTCTTAATGAGAGCTTTTGTGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT 611
 605 TCTCTTAATGAGAGCTTTTGTGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT 664
 612 GATAAATGAGAGTGTGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT 671
 665 GATAAATGAGAGTGTGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT 724
 672 TTTTCTTCAACGACGTTTCTGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT 731
 725 TTTTCTTCAACGACGTTTCTGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT 784
 732 TGTGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT 784
 785 TGTGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT 844
 785 ----- 784
 845 GAATTAATAATAAAGTACGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT 904
 785 ----- 784

PD 2/19/00

2

Db 905 TCACATGATGCGAGATGCTAGTCTCAGGCTCAGTGATAGTCTGTATATATATGAT 964
 785 ----- 784
 Db 965 ACTAATACAGATATATATCTTCAACATTTGACTCAGACACAGATATGTCACACTTGT 1024
 785 ----- 784
 Db 1025 GCTTTTGACCTAATACCTTTTACTGTCATGTTCAATGACAAAACAGTGAACATC 1084
 785 ----- 784
 Db 1085 TGGCAATTTGACCTGGAACACTTTTGCAGCAGAGACACAGAACATGCTGAGACAA 1144
 816 TTTACGGAATTTGTCAGAGAGTGTCTCAACATGCTTTGTGTCACAAAGATTTAAA 875
 1145 TTTACGGAATTTGTCAGAGAGTGTCTCAACATGCTTTGTGTCACAAAGATTTAAA 1204
 876 GATCTGTTGATTTTTCAGATGATGATTAACATTTGATGTAAGAACTGTTGAATCTTACA 935
 1205 GATCTGTTGATTTTTCAGATGATGATTAACATTTGATGTAAGAACTGTTGAATCTTACA 1264
 936 AAGAAAGTCTGCTGATGATTTGAAATTTGATGATGTCGATGTCGATGTCGATGTCGATGTCGAT 995
 1265 AAGAAAGTCTGCTGATGATTTGAAATTTGATGATGTCGATGTCGATGTCGATGTCGATGTCGAT 1324
 996 AAGAAATTTGAAGCTCAGACACCAAGTTAAATCCCTTTCTCAGGAATTCGATGAA 1055
 1325 AAGAAATTTGAAGCTCAGACACCAAGTTAAATCCCTTTCTCAGGAATTCGATGAA 1384
 1056 TTTATATGTCATTAATCTAGAGAACTTATGAAGATCCGTCATGTCATGATGTCAT 1115
 1385 TTTATATGTCATTAATCTAGAGAACTTATGAAGATCCGTCATGTCATGATGTCAT 1444
 1116 TCATATGAAAAGAGCAATGTAAGATTTGATGACGACAAAAGAAAGCTACAGTCCATG 1175
 1445 TCATATGAAAAGAGCAATGTAAGATTTGATGACGACAAAAGAAAGCTACAGTCCATG 1504
 1176 ACAATATTTCTTCTCTTACGAGTACCTTACACCAATTTGATGACGACAAAAGAAAGCTACAGTCCATG 1235
 1505 ACAATATTTCTTCTCTTACGAGTACCTTACACCAATTTGATGACGACAAAAGAAAGCTACAGTCCATG 1564
 1236 AATGATGCTGAGACACACCAAAAGTAA 1266
 1565 AATGATGCTGAGACACACCAAAAGTAA 1595

RESULT 3
 CQ729621 1254 bp DNA linear PAT 03-FEB-2004
 DEFINITION Sequence 15555 from Patent WO02068579.
 ACCESSION CQ729621
 VERSION CQ729621.1 GI:42301208
 KEYWORDS

SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
 Kils, such as nucleic acid arrays, comprising a majority of
 humanexons or transcripts, for detecting expression and other uses
 thereof
 Patent: WO 02068579-A 15555 06-SEP-2002;
 JOURNAL
 PR Corporation (NY) (US)
 Location/Qualifiers
 1.1254
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

FEATURES
 SOURCE

ORIGIN

Query Match	71.7%	Score 912.6	DB 6	Length 1254									
Best Local Similarity	86.5%	Pred. No. 2.5e-236											
Matches 1092	Conservative 0	Mismatches 54	Indels 117	Gaps 3									
QY	111	ATGTGAAA	CTGATT	CACACATT	AGCTGAT	TCATGTG	TGACGAT	GTGAC	GTGTC	CTTC	170		
Db	1	ATGTGAAA	CTGATT	CACACATT	AGCTGAT	TCATGTG	TGACGAT	GTGAC	GTGTC	CTTC	60		
QY	121	TCCTTTT	CCCTCT	TGGCTA	CTTGCT	TCCTT	GACAAA	CAATT	CGCCT	GTA	CTGTTA	230	
Db	61	TCCTTTT	CCCTCT	TGGCTA	CTTGCT	TCCTT	GACAAA	CAATT	CGCCT	GTA	CTGTTA	120	
QY	231	GACCTT	ACGAA	CTGCAC	ATTCT	CCATT	GAA	GTTCAT	TA	CTT	AGCTG	290	
Db	121	GACCTT	ACGAA	CTGCAC	ATTCT	CCATT	GAA	GTTCAT	TA	CTT	AGCTG	180	
QY	291	TGTTT	CTCCCT	TCAGCA	CATAT	TTT	TGGCAT	TGTT	GAACA	AGAT	GTACCA	CTGCTA	350
Db	181	TGTTT	CTCCCT	TCAGCA	CATAT	TTT	TGGCAT	TGTT	GAACA	AGAT	GTACCA	CTGCTA	240
QY	351	TGGAAT	ACTGAA	AATGGA	CAGAT	GTGCG	ATGAT	GAA	CAGCCT	TA	GTGG	CA	410
Db	241	TGGAAT	ACTGAA	AATGGA	CAGAT	GTGCG	ATGAT	GAA	CAGCCT	TA	GTGG	CA	300
QY	411	AGGGTT	TGCA	GTCTT	TCCCA	CACT	CCAC	GTGTT	TGGCAT	CA	GGGGG	CA	470
Db	301	AGGGTT	TGCA	GTCTT	TCCCA	CACT	CCAC	GTGTT	TGGCAT	CA	GGGGG	CA	360
QY	471	GTCGTTT	TGTGGA	ATGCA	CAGT	CAT	CA	TAA	TTAT	TA	TAGAT	GTGAT	530
Db	361	GTCGTTT	TGTGGA	ATGCA	CAGT	CAT	CA	TAA	TTAT	TA	TAGAT	GTGAT	420
QY	531	TCCTTTG	GGGCA	TGTG	CATTTT	CTCCT	TA	TGGA	AGCTT	CTTT	G	CA	590
Db	421	TCCTTTG	GGGCA	TGTG	CATTTT	CTCCT	TA	TGGA	AGCTT	CTTT	G	CA	480
QY	591	GSTGAT	TTTA	CAGT	GTGG	GA	TGAT	TA	AAAT	GAGT	GT	CT	650
Db	481	GSTGAT	TTTA	CAGT	GTGG	GA	TGAT	TA	AAAT	GAGT	GT	CT	540
QY	651	CTTGAAT	TTAC	CTGT	CTGG	ATTTT	CTT	CA	CAGC	AGTT	CT	GAT	704
Db	541	CTTGAAT	TTAC	CTGT	CTGG	ATTTT	CTT	CA	CAGC	AGTT	CT	GAT	600
QY	705	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	704	
Db	601	AAAAGT	AC	TGAG	TGGG	CA	CTGT	CT	CTGT	CT	GTG	CT	660
QY	705	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	705	
Db	661	CAGAT	G	TAGT	CT	CAGG	AT	AT	GT	CA	CA	CT	720
QY	726	GCAT	CAT	GT	GTG	GT	GC	AA	GT	GC	AA	AT	782
Db	721	GCAT	CAT	GT	GTG	GT	GC	AA	GT	GC	AA	AT	780
QY	783	TTAG	CAG	GG	CG	CA	GAA	CAT	CA	GT	CA	GT	842
Db	781	CAAG	CA	AG	GA	GA	CA	CT	AG	TT	TA	CG	840
QY	843	GT	CT	CA	CA	TG	CT	TT	GT	CA	CA	AG	902
Db	841	GT	CT	CA	CA	TG	CT	TT	GT	CA	CA	AG	900
QY	903	AA	CAT	TG	AT	G	AA	AA	AA	AA	AA	AA	962
Db	901	AA	CAT	TG	AT	G	AA	AA	AA	AA	AA	AA	960
QY	963	ATT	GAT	CT	T	A	G	A	CT	G	A	T	1022
Db	961	ATT	-----	-----	-----	-----	-----	-----	-----	-----	-----	1011	
QY	1023	GT	TAA	AT	CC	T	T	T	T	T	T	T	1082

D _b	1012	GTTAAATCCCTTTCCTTAGGAATTCCTGATGMAATTATATGTCCAAATACCTAGGAACTT	1071
O _y	1083	ATGAAAGATCCGGTCATCGCATCAGATGGCTATTCAATGAAAAGAAAGCAATGGAAAT	1142
D _b	1072	ATGAAAGATCCGGTCATCGCATCAGATGGCTATTCAATGAAAAGAAAGCAATGGAAAT	1131
O _y	1143	TGATCAGCAAAAAAGAAACGTACAAGTCCCATAGCMAATCTTGTTCTTCTTCACGGTA	1202
D _b	1132	TGATCAGCAAAAAAGAAACGTACAAGTCCCATAGCMAATCTTGTTCTTCTTCACGGTA	1191
O _y	1203	CTTACACCAATATAGACTCTGMAAATGGCCATCAATGAGCGTGGAGACACACCAAAAG	1262
D _b	1192	CTTACACCAATATAGACTCTGMAAATGGCCATCAATGAGCGTGGAGACACACCAAAAG	1251
O _y	1263	TAA 1265 	
D _b	1252	TAA 1254 	

	RESULT	4
	LOCUS	AKI29983
	DEFINITION	AKI29983 1919 bp mRNA linear PRI 10-SEP-2003
	ACCESSION	AKI29983
	VERSION	AKI29983.1 GI:34526702
	KEYWORDS	cddo capping; fts (full insert sequence).
	SOURCE	Homo sapiens (human)
	ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
	REFERENCE	1 Teshiro, H., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Suzuki, Y., Hata, H., Nakagawa, K., Mizuno, S., Morinaga, M., Kawamura, M., Sugiyama, T., Iite, R., Otsuki, T., Sato, H., Nishikawa, T., Sugiyama, A., Kawakami, B., Nagai, K., Isogai, T. and Sugano, S. NEDO human cDNA sequencing project Unpublished
TITLE	JOURNAL	2 (bases 1 to 1919) Sugano, S. and Suzuki, Y. Direct Submission Submitted (31-JUL-2003) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; Shitake-da, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail: litchda@ns.u-tokyo.ac.jp, Tel: 81-3-5449-5286, Fax: 81-3-5449-5416)
COMMENT		NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction and 5'-end one pass sequencing: Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; 3'-end one pass sequencing: RAB; clone selection for full insert sequencing: RAB and Helix Research Institute. Location/Qualifiers 1..1919 /organism="Homo sapiens" /mol_type="RNA" /db_xref="taxon:9606" /cldne="KDN04521" /tissue_type="kidney" /clone_lib="KDN" /note="Cloning vector: pME18sFL3"
FEATURES	source	
	ORIGIN	
Query Match		70.2%; Score 893.4; DB 9; Length 1919;
Best Local Similarity		80.8%; Pred. No. 4.3e-231;
Matches 1183; Conservative		0; Mismatches 6; Indels 276; Gaps 1,
OY		78 CTCCTGAGGCCTTTTTCCTCAATTAAGAACAATGTGAACATTGACTTGACATACTTAGCT 137
Db		225 CCCTGGCGGGCTGTTTTTCTTCAAATAAAMAACTGGTGAACGATTCCACACATTAGCT 284

OY	138	NATCATGGTGAAGATGCAACGCGATGACCTTCTCCCTTCCCTTGGCACTATGGCC	197
Db	285	GATCATGGTGAAGATGCAACGCGATGACCTTCTCCCTTGGCACTATGGCC	344
OY	198	TTGGACAAMAAATTCGCTGTACTCGTTAAGTGACTTTACTGAATCCCACTTCTCCA	257
Db	345	TTGGACAAMAAATTCGCTGTACTCGTTAAGTGACTTTACTGAATCCCACTTCTCCA	404
OY	258	TTGAAGTTTCATACCTATGCTGTCCAGCTGCTGTCTTCTCCCTTCAGACATATTTTG	317
Db	405	TTGAAGTTTCATACCTATGCTGTCCAGCTGCTGTCTTCTCCCTTCAGACATATTTTG	464
OY	318	GCATCGTGTCAACAGATGTACCACTGTCCATAGAAATACCTGAAMAAATGACAGATGCTG	377
Db	465	GCATCGTGTCAACAGATGTACCACTGTCCATAGAAATACCTGAAMAAATGACAGATGCTG	524
OY	378	GCAGTATGTAAACAGCTAGTGGCAGCCCTGTGAGGGTTTGGCAGTTTTCACAGACTCC	437
Db	525	GCAGTATGTAAACAGCTAGTGGCAGCCCTGTGAGGGTTTGGCAGTTTTCACAGACTCC	584
OY	438	ACGTGTTTGGCATCAGGGGACGCTGATGAACTGTGTTTTGTGTGAATGACACAGTCAATAC	497
Db	585	ACGTGTTTGGCATCAGGGGACGCTGATGAACTGTGTTTTGTGTGAATGACACAGTCAATAC	644
OY	498	AAATTAATTAATGTGTGTATGTATTAATATGGCTCTTGGGGGCAATGTGCATTTTCTCTCT	557
Db	645	AAATTAATTAATGTGTGTATGTATTAATATGGCTCTTGGGGGCAATGTGCATTTTCTCTCT	704
OY	558	AATGGAACCTTCTTGTCACTGGGCTCTCATGTGTGTATTAACAGTGTGGATGTATAA	617
Db	705	AATGGAACCTTCTTGTCACTGGGCTCTCATGTGTGTATTAACAGTGTGGATGTATAA	764
OY	618	ATGAGGTGTCTGATAGTAAAMAGACATGATCTTGGAAATTAACCTGTGTGCATTTTCT	677
Db	765	ATGAGGTGTCTGATAGTAAAMAGACATGATCTTGGAAATTAACCTGTGTGCATTTTCT	824
OY	678	TCACAGCCAGTTTCTGATGAGAGAACAGAGTCTTCAATTTTTTGCATGTGCATCATGTGT	737
Db	825	TCACAGCCAGTTTCTGATGAGAGAACAGAGTCTTCAATTTTTTGCATGTGCATCATGTGT	884
OY	738	CAGATTTGCCAAGTCAMAAATTTGATGTGTTCTTTTACCATATCTT-----	784
Db	885	CAGATTTGCCAAGTCAMAAATTTGATGTGTTCTTTTACCATATCTTAGCTTTGAATTA	944
OY	785	-----	784
Db	945	AAATTAATAAGTACGTAGTGGGCACTGTGCTCTGTGTTGGCTTGTGCTTTTCCCAT	1004
OY	785	-----	784
Db	1005	GATGGCAGATGCTAGTCTCAGGGTCAATGATATAGTGTATATGATATATAT	1064
OY	785	-----	784
Db	1065	ACTGAGAAATATCTTACACATGTACTGACGACACACAGGATATGCAACACTTGTCCTTTT	1124
OY	785	-----	784
Db	1125	GCACCTAATACCTTTTACTTGTCTACTGTGTTCAATGAGCAAMAAACGTGAACATCTGGCAA	1184
OY	785	-----AGCAAGCGCACAGAACATCAGCTGAGCAATTTTACC	821
Db	1185	TTTGAAGCTGAGAAACATTTTGCACCAAGAGACAGAAACATCAGCTGAGCAATTTTACC	1244
OY	822	GAAATTTGTGTGAGAGAGTGTCTCAACATGGCTTTGTGCAACAGATTTTAAAGATCTT	881
Db	1245	GAAATTTGTGTGAGAGAGTGTCTCAACATGGCTTTGTGCAACAGATTTTAAAGATCTT	1304
OY	882	GTTGGTATTTTCAAGATGATTAACATTTGATGTGAAAAAGAACTGTGAACTTTTCAAAAGAA	941
Db	1305	GTTGGTATTTTCAAGATGATTAACATTTGATGTGAAAAAGAACTGTGAACTTTTCAAAAGAA	1366
OY	942	AGTCTGGCTGATGATTTTGAAMATGAAATCTTAGAGCTGCGTAGTAAGTCTGAGAAA	1001

Db	1365	AGTGGCTGATGATTTGAAAATTGGAATCTCTAGAGCTGGCTAGTAAAGCTGAGGAAA	1424
Qy	1002	ATTGAGAGGCTCAGAGCAAGGTTAAATCCCTTTCTTCAGAAATTCCTGATGAAATTTATA	1061
Db	1425	ATTGAGAGGCTCAGAGCAAGGTTAAATCCCTTTCTTCAGAAATTCCTGATGAAATTTATA	1484
Qy	1062	TGTCCTAATTAATGAGAACTTATGAAAAGATCCGGTCACTGGCTATGATGCGCTATTCTAT	1121
Db	1485	TGTCCTAATTAATGAGAACTTATGAAAAGATCCGGTCACTGGCTATGATGCGCTATTCTAT	1544
Qy	1122	GAAAAGGAGCAATGAGAAAATTGGATCAGCAAAAAGAAAGATACAGCTCCCATGACAAAT	1181
Db	1545	GAAAAGGAGCAATGAGAAAATTGGATCAGCAAAAAGAAAGATACAGCTCCCATGACAAAT	1604
Db	1605	CTTGCTTCTCTCTTCAGCGGATCTTACACCAAAATGAGACTGAAAATGGCCATCAATAGA	1684
Qy	1242	TGCGTCGAGACACACCCAAAAGTAAA	1266
Db	1665	TGCGTCGAGACACACCCAAAAGTAAA	1689
RESULT 5			
AX748164			
LOCUS	AX748164	1811 bp	mRNA
DEFINITION	Sequence 1689 from Patent Epi308459.		linear
ACCESSION	AX748164		
VERSION	AX748164.1	GI:32132552	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE			
AUTHORS	1 Isegai,T., Sugiyama,T., Otsuki,T., Makamatsu,A., Sato,H., Ishii,S., Yamamoto,J.T., Isono,Y., Hio,Y., Otsuka,K., Nagai,K., Irie,R., Tamechika,I., Seki,N., Yoshikawa,T., Otsuka,M., Nagahari,K. and Maezho,Y.		
TITLE	Full-length cDNA sequences		
JOURNAL	Patent: EP 1308459-A 1689 07-MAY-2003;		
	Helix Research Institute (UP) ; Research Association for Biotechnology (UP)		
FEATURES			
source	location/Qualifiers		
	1..1811		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match	70.0%; Score 890.2; DB 6; Length 1811;		
Best Local Similarity	80.9%; Pred. No. 3,2e-230;		
Matches 1178; Conservative	0; Mismatches 3; Indels 276; Gaps 1;		
Qy	86	GGCTGTTTCTTCAATAAAGAACATGTGAAACGATTCACACATTAAGCTGATCATGG	145
Db	125	GGCTGTTTCTTCAATAAAGAACATGTGAAACGATTCACACATTAAGCTGATCATGG	184
Qy	146	TGAAGAGTCAACGTGCTGTGCTTCTCTTTCCCTTGGCTACTTGCTCTTGACAA	205
Db	185	TGACGAGTCAACGTGCTGTGCTTCTCTTTCCCTTGGCTACTTGCTCTTGACAA	244
Qy	206	AACAAATCGCCTGACTCGTTACGTGACTTAACTGAACTGCACTTCATTAAGATT	265
Db	245	AACAAATCGCCTGACTCGTTACGTGACTTAACTGAACTGCACTTCATTAAGATT	304
Qy	266	TCATACCTATGCTGTCAGCTGCTGTGTTCTCCCTTCAGACATAATTTGGCATCGT	325
Db	305	TCATACCTATGCTGTCAGCTGCTGTGTTCTCCCTTCAGACATAATTTGGCATCGT	364
Qy	326	TTCAACAGATGTAACCACTGTCTTATGGAATCTGAAAATGACATGATGCTGGAGAT	385

Db 365 TTCACAGATGATGACCACTGTCTTATGGAATACTGAAATATGACAGATGCTGGCAGATGAT 424
 QY 386 GGAACAGCCTAGTGGAGCCCTGTGAGAGGGTTTGGCCAGATTTCGCCAGACTCCAGCTGTTT 445
 Db 425 GGAACAGCCTAGTGGAGCCCTGTGAGAGGGTTTGGCCAGATTTCGCCAGACTCCAGCTGTTT 484
 QY 446 GGCATCAGGGGAGCTGATGGAACCTGTGTTTGTGGAAATGACAGTCAATCAAAATTATA 505
 Db 485 GGCATCAGGGGAGCTGATGGAACCTGTGTTTGTGGAAATGACAGTCAATCAAAATTATA 544
 QY 506 TAAATGTGTAAGTGTAAAGATAGGCTCTCTGGGGGAGCTGTCATTTTCTCTTAATGGAAG 565
 Db 545 TAAATGTGTAAGTGTAAAGATAGGCTCTCTGGGGGAGCTGTCATTTTCTCTTAATGGAAG 604
 QY 566 CTCTCTTGTGATCAGTGGCTCTCATGTGGTGTATTTAAACAGTGTGGAGATGATAAATGAGGTG 625
 Db 605 CTCTCTTGTGATCAGTGGCTCTCATGTGGTGTATTTAAACAGTGTGGAGATGATAAATGAGGTG 664
 QY 626 TCTGCATAGTGAAGAAAGACATGATCTTGGAAATTAACCTGTGGATTTTCTTCACAGCC 685
 Db 665 TCTGCATAGTGAAGAAAGACATGATCTTGGAAATTAACCTGTGGATTTTCTTCACAGCC 724
 QY 686 AGTTTGTGATGGAACCAAGGTCTTCAGTGTTCCTTTCAGCTGGCATCATGTGTGACGATTTG 745
 Db 725 AGTTTGTGATGGAACCAAGGTCTTCAGTGTTCCTTTCAGCTGGCATCATGTGTGACGATTTG 784
 QY 746 CCAAGTCAAAATTTGGATTGTTCTTTTACCCTATCTT----- 784
 Db 785 CCAAGTCAAAATTTGGATTGTTCTTTTACCCTATCTT----- 844
 QY 785 ----- 784
 Db 845 AAGTACACTGATGAGGCACTGTCTCTGTTCTTGCTTGTGCTTTTCCCATGATGGGCA 904
 QY 785 ----- 784
 Db 905 GATGCTAGTCTCAGGGTCACTGATGATGCTGTCTATGATATGATATTAATCTAGTGAAG 964
 QY 785 ----- 784
 Db 965 TATATCTTCAACATTTGACTGACACACACAGGTATGTCAACAATTTGCTTTTGACCTTA 1024
 QY 785 ----- 784
 Db 1025 TACCTTTTACTGTTACTGTTCAATGAGCAAAACAGTGAACATCTGGCAATTTGACT 1084
 QY 785 -----AGCAAGGGGCAACAAACATCAGCTGAACCAATTTACCGAAGATTG 829
 Db 1085 GGAACACCTTTGGCCAAAGGAGGACACAGAACATCAGCTGAACCAATTTACCGAAGATTG 1144
 QY 830 GTGAGAGAGAGTGTCTCAACATGAGCTTTGTGCAAAAGATTTTAAAGATCTTGTGGTAT 889
 Db 1145 GTGAGAGAGAGTGTCTCAACATGAGCTTTGTGCAAAAGATTTTAAAGATCTTGTGGTAT 1204
 QY 890 TTTCAAGATGATTAACATTTGATGAGAAAGAACTGTTGAATCTTTACAAAGAAAGTCTGGC 949
 Db 1205 TTTCAAGATGATTAACATTTGATGAGAAAGAACTGTTGAATCTTTACAAAGAAAGTCTGGC 1264
 QY 950 TGAATGATTTGAAAATTTGAATCTCTAGAGACTGCGTAGTAAGTCTGAGGAAATTTGAAGA 1009
 Db 1265 TGAATGATTTGAAAATTTGAATCTCTAGAGACTGCGTAGTAAGTCTGAGGAAATTTGAAGA 1324
 QY 1010 GCTCAGGAGCAAGGTTAAATCCCTTTCTTCAAGGAATCTGATGAATTTTATGTCCAAAT 1069
 Db 1325 GCTCAGGAGCAAGGTTAAATCCCTTTCTTCAAGGAATCTGATGAATTTTATGTCCAAAT 1384
 QY 1070 AACTAGAGAACTTATGAAAGATCCGGTCAATCGATCGATCGATGCGTATTCATATGAAAAGA 1129
 Db 1385 AACTAGAGAACTTATGAAAGATCCGGTCAATCGATCGATCGATGCGTATTCATATGAAAAGA 1444
 QY 1130 AGCAATGGAATAATTTGATCAGCAAAAAGAAAGCTAAGTCCCATGCAAAATCTTGTCTT 1189
 Db 1445 AGCAATGGAATAATTTGATCAGCAAAAAGAAAGCTAAGTCCCATGCAAAATCTTGTCTT 1504

QY 1190 TCCTTCAGCGGTACTTACACCAAAATGAGACTCTGAAAATGCGCATCAATAGATGCGTGA 1249
 Db 1505 TCCTTCAGCGGTACTTACACCAAAATGAGACTCTGAAAATGCGCATCAATAGATGCGTGA 1564
 QY 1250 GACACACCAAAAGTAAA 1266
 Db 1565 GACACACCAAAAGTAAA 1581
 RESULT 6
 AK093494
 LOCUS
 DEFINITION
 Homo sapiens cDNA FLJ36175 f18, clone TEST12026453, weakly similar to PUTATIVE SERINE/THREONINE-PROTEIN KINASE PKWA (EC 2.7.1.-).
 ACCESSION
 AK093494
 VERSION
 AK093494.1 GI:21752384
 KEYWORDS
 Oligo capping; f18 (full insert sequence).
 SOURCE
 Homo sapiens
 Homo sapiens (human)
 ORGANISM
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 1
 Oca.T., Suzuki.Y., Nishikawa.T., Otsuki.T., Sugiyama.T., Irie.R.,
 Wakamatsu.A., Hayashi.K., Sato.H., Nagai.K., Kimura.K., Makita.H.,
 Sekine.M., Obauchi.M., Nishi.T., Shibahara.T., Tanaka.T.,
 Ishii.S., Yamamoto.J., Saito.K., Kawai.Y., Isono.Y., Nakamura.Y.,
 Nagahari.K., Murakami.K., Yasuda.T., Iwayanagi.T., Wagatsuma.M.,
 Shiratori.A., Sudo.H., Hosokiri.T., Kaku.Y., Kodaira.H., Kondo.H.,
 Sugawara.M., Takahashi.M., Kanda.K., Yokoi.T., Furuya.T.,
 Kikawa.E., Omura.Y., Abe.K., Kamihara.K., Katsuta.N., Sato.K.,
 Tanikawa.M., Yamazaki.M., Minomura.K., Ishida.S., Yamashita.H.,
 Muraoka.K., Fujimori.K., Tanai.H., Kimata.M., Watanabe.M.,
 Hirooka.S., Chiba.Y., Ichida.S., Ono.Y., Takiguchi.S., Watanabe.S.,
 Yoshida.M., Hoshino.T., Kusano.J., Kanehori.K., Takahashi-Fujii.A.,
 Hara.H., Tanabe.T., Nomura.Y., Togiyama.S., Komai.F., Hara.R.,
 Takeuchi.K., Arita.M., Imose.N., Musashino.K., Yuuki.H., Oshima.A.,
 Sasaki.N., Aotsuka.S., Yoshikawa.Y., Matsunawa.H., Ichihara.T.,
 Shiohara.N., Sano.S., Moriya.S., Momiyama.H., Satoh.N., Takami.S.,
 Terehima.Y., Suzuki.O., Nakagawa.S., Senoh.A., Mizoguchi.H.,
 Goto.Y., Shimizu.F., Wakebe.H., Hishigaki.H., Watanabe.T.,
 Sugiyama.A., Takemoto.M., Kawakami.B., Yamazaki.M., Watanabe.K.,
 Kumagai.A., Itakura.S., Fukuzumi.Y., Fujimori.Y., Komiyama.M.,
 Tashiro.H., Tanigami.A., Fujiwara.T., Ono.T., Yamada.K., Fujii.Y.,
 Ozaki.K., Hirao.M., Ohmori.Y., Kawabara.A., Hikiji.T., Kobatake.N.,
 Inagaki.H., Ikema.Y., Okamoto.S., Ohtani.R., Kawakami.T.,
 Noguchi.S., Itoh.T., Shigeta.K., Senba.T., Matsumura.K.,
 Nakajima.Y., Mizuno.T., Morinaga.M., Sasaki.M., Togashi.T.,
 Oyama.M., Hata.H., Watanabe.M., Komatsu.T., Mizushima-Sugano.J.,
 Satoh.T., Shirai.Y., Takahashi.Y., Nakagawa.K., Okumura.K.,
 Nagase.T., Nomura.N., Kikuchi.H., Masuno.Y., Yamashita.R.,
 Nakai.K., Yada.T., Nakamura.Y., Ohara.O., Isogai.T. and Sugano.S.
 Complete sequencing and characterization of 21,243 full-length
 human cDNAs
 JOURNAL
 PUBMED
 14702039
 Nat. Genet. 36 (1), 40-45 (2004)
 2
 Ishibashi.T., Kanehori.K., Yoshida.M., Watanabe.S., Ishida.S.,
 Ono.Y., Hoshino.T., Hirooka.S., Murakami.K., Takiguchi.S.,
 Kusano.J., Watanabe.M., Fujimori.K., Tanai.H., Ishida.M.,
 Yamashita.H., Chiba.Y., Sugiyama.T., Irie.R., Otsuki.T., Sato.H.,
 Wakamatsu.A., Ishii.S., Yamamoto.J., Isono.Y., Kawai-Hio.Y.,
 Saito.K., Nishikawa.T., Kimura.K., Matsuno.K., Nakamura.Y.,
 Sekine.M., Kikuchi.H., Kanda.K., Wagatsuma.M., Takahashi-Fujii.A.,
 Oshima.A., Sugiyama.A., Kawakami.B., Suzuki.Y., Sugano.S.,
 Nagahari.K., Masuno.Y., Nagai.K. and Isogai.T.
 NEDO human cDNA sequencing project
 Unpublished
 3 (bases 1 to 1811)
 Isogai.T. and Yamamoto.J.
 Direct Submission
 Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team) ; 2-6-7
 Kazuo-Kamatari, Kiseazaki, Chiba 292-0812, Japan

COMMENT

(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
 NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

FEATURES

Source

1. 1811
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="TEST12026453"
 /issue_type="testis"
 /clone_1db="TEST12"
 /note="Cloning vector: pME18SFL3"
 150. 1580
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BAC04184.1"
 /db_xref="GI:21752385"
 /translation="MVKLHTLADHGDVNCASFSLATCSLDKTIKLYSLDPTF
 LPHSLKFTHTYAVHCCPSPSGHILASCTDGTIVLWNTENGOMALPSPNPSFVSV
 GQSPDSTCLASGADGTIVLWNAOSYLYRGSGVKGSLAASPSPSPVSSC
 GDLVMDKMRCLSEKADHGLTICDPSQPVSDGQLOFPLASGDCQVYKIN
 VSFTHILGFLKTKSTLSGHCAPYLACFSHDGOMLVSGSDVIVYDINTEIHLT
 LTHRTYVTTCAFPNVLTLNAGSMKRTVNIWQDLDTLCQARSIEHQLKPTDWSR
 EDVSTWLCADQLKLVGIFPKNNIDGKELMLTKESLADLKISLIGRSVLYKIEE
 LRTKYSLSGIPDEFICPIRELKMDPVIASDYSYEKEMENWISKKRTSPMTNL
 VLPSAVLTPLNRTLMALINRMLETQHK"

CDS

ORIGIN

Query Match 70.0%; Score 890.2; DB 9; Length 1811;
 Best Local Similarity 80.9%; Pred. No. 3.2e-230;
 Matches 1178; Conservative 0; Mismatches 3; Indels 276; Gaps 1;

86 GGCCTGTTTCTTCAATAAAGACATGTAACCTGTAACATTAAGCTGATCAGG 145
 125 GCGTGTCTTCAATAAAGACATGTAACCTGTAACATTAAGCTGATCAGG 184
 146 TGACGATGTCAACTGCTGCTCTCTCTTCCCTTGGCTACTTCTCTTGGACAA 205
 185 TGACGATGTCAACTGCTGCTCTCTCTTCCCTTGGCTACTTCTCTTGGACAA 244
 206 AACATTCGCTGTAAGTACGTAACCTGTAACCTGTAACCTGTAACCTGTA 265
 245 AACATTCGCTGTAAGTACGTAACCTGTAACCTGTAACCTGTAACCTGTA 304
 266 TCATACCTGTAAGTACGTAACCTGTAACCTGTAACCTGTAACCTGTAACCT 325
 305 TCATACCTGTAAGTACGTAACCTGTAACCTGTAACCTGTAACCTGTAACCT 364
 326 TTCAACGATGTCAACTGCTGCTCTCTCTTCCCTTGGCTACTTCTCTTGGAC 385
 365 TTCAACGATGTCAACTGCTGCTCTCTCTTCCCTTGGCTACTTCTCTTGGAC 424
 386 GGAACAGCTAGTGGACGCTGTGAGGGTTCAGAGTTTCCCAAGCTCAGCTGTT 445
 425 GGAACAGCTAGTGGACGCTGTGAGGGTTCAGAGTTTCCCAAGCTCAGCTGTT 484
 446 GGCATCAGGGGACGCTGATGGAACCTGTGTTTGTGGAATGACAGTATCAATTA 505
 485 GGCATCAGGGGACGCTGATGGAACCTGTGTTTGTGGAATGACAGTATCAATTA 544
 506 TAGATGTGATGTGTTAAAGATGCTCTTGGCGGACATGATCTTCTCTTAATGA 565
 545 TAGATGTGATGTGTTAAAGATGCTCTTGGCGGACATGATCTTCTCTTAATGA 604
 566 CTCTCTTGTCACTGGCTCTTCAATGTGATGTTAAACAGTGTGGATGATTAATG 625
 605 CTCTCTTGTCACTGGCTCTTCAATGTGATGTTAAACAGTGTGGATGATTAATG 664

626 TCTGCATAGTGAAGAAACACATGATCTTGGAAATTAACCTGCTGCGATTTTCTTCA 685
 665 TCTGCATAGTGAAGAAACACATGATCTTGGAAATTAACCTGCTGCGATTTTCTTCA 724
 686 AGTTTCTGATGAGAACAGAGCTTTCAGTTTTCAGCTGATCATGATGTCAGATG 745
 725 AGTTTCTGATGAGAACAGAGCTTTCAGTTTTCAGCTGATCATGATGTCAGATG 784
 746 CCAAGTCAAAATTTGATTTGTTTCTTCAATCTT----- 784
 785 CCAAGTCAAAATTTGATTTGTTTCTTCAATCTTAGTTTGAATTAATTA 844
 785 ----- 784
 845 AAGTACACTGAGTGGACATGCTCTCTGTTCTGCTTTCCTTCCCATGATGAGCA 904
 785 ----- 784
 905 GATGCTAGTCTCAGGGTCAAGTATAGTGTATGATGATGATTAATTAATGAGAA 964
 785 ----- 784
 965 TATACCTCAGACATTTGATCTGACACACAGGATATGCAAACTTGCTTTTGACCTAA 1024
 785 ----- 784
 1025 TACCCTTTACTTCTGCTGCTGCTTCAATGACAAACAGTGAACATCTGGCAATTTG 1084
 785 -----ACCAAGCCGACAGAACATCAGCTGAGAACATTTACCGAAGATG 829
 1085 GGAACACTTTGCCAACAGAGACAGAACATCAGCTGAGAACATTTACCGAAGATG 1144
 830 GTCAAGAGAGTGTCTTCAACATGCTTTTGTGCAAGATTTAAAGATCTTGTGAT 889
 1145 GTCAAGAGAGTGTCTTCAACATGCTTTTGTGCAAGATTTAAAGATCTTGTGAT 1204
 890 TTTCAAGATGATTAACATTTGATGAGAAAGAGTGTGATCTTCAAAAGAGTCTGGC 949
 1205 TTTCAAGATGATTAACATTTGATGAGAAAGAGTGTGATCTTCAAAAGAGTCTGGC 1264
 950 TGATGATTTGAAATTAATCTCTGAGATGCTGATTAAGTGTGAGAAATTTGAGA 1009
 1265 TGATGATTTGAAATTAATCTCTGAGATGCTGATTAAGTGTGAGAAATTTGAGA 1324
 1010 GCTCAGACCAAGTTAAATCCTTTCTTGAAGATTTCTGATGATTTATATGTCAT 1069
 1325 GCTCAGACCAAGTTAAATCCTTTCTTGAAGATTTCTGATGATTTATATGTCAT 1384
 1070 AACTAGAGACTTATGAAAGATCCGTCATGCTGATGATGCTGATGATGATGAA 1129
 1385 AACTAGAGACTTATGAAAGATCCGTCATGCTGATGATGCTGATGATGATGAA 1444
 1130 AGCAATGGAATAATGATCAGCAAAAGAAAGTCAAGTCCCATGCAAAATCTTGTCT 1189
 1445 AGCAATGGAATAATGATCAGCAAAAGAAAGTCAAGTCCCATGCAAAATCTTGTCT 1504
 1190 TCTTCAAGCGTACTTACACCAATATGACTCTGAAATGCGCATCAATGATGCTG 1249
 1505 TCTTCAAGCGTACTTACACCAATATGACTCTGAAATGCGCATCAATGATGCTG 1564
 1250 GACACACCAAAAGTAA 1266
 1565 GACACACCAAAAGTAA 1581

RESULT 7
 LOCUS CQ491312 1596 bp DNA linear PAT 30-JAN-2004
 DEFINITION Sequence 23179 from Patent WO0160860
 ACCESSION CQ491312
 VERSION CQ491312.1 GI:41456931
 KEYWORDS

SOURCE	ORGANISM	Homosapiens (human)
REFERENCE	AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	TITLE	Schlegel, R., Endege, W.O. and Monahan, J.E. Genes differentially expressed in human prostate cancer and their use
JOURNAL	Patent: WO 0160860-A 23179 23-AUG-2001; Millennium Predictive Medicine, Inc. (US)	
FEATURES	source	Location/Qualifiers 1..1996 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"
ORIGIN		
Query Match	69.1%;	Score 878.8; DB 6; Length 1996;
Best Local Similarity	80.7%;	Pred. No. 4e-227;
Matches 1166;	Conservative 0;	Mismatches 2; Indels 276; Gaps 1;
QY	99	CAATAAAGAATGATGAAATCTGATTCACACATTAGTATCATGTGACGATGTCAAC 158
DB	4	CAATAAAGAATGATGAAATCTGATTCACACATTAGTATCATGTGACGATGTCAAC 63
QY	159	TGCTGTGCTCTCTCTTTTCCCTCTTGAGTACTGCTGCTTGAAGAAATTCGCTG 218
DB	64	TGCTGTGCTCTCTCTTTTCCCTCTTGAGTACTGCTGCTTGAAGAAATTCGCTG 123
QY	219	TACTGTAGTACGATCTTACTGAACTGACACATTCCTCATGTAATTCATACCTACT 278
DB	124	TACTGTAGTACGATCTTACTGAACTGACACATTCCTCATGTAATTCATACCTACT 183
QY	279	GTCCACTGCTGCTGTTTCTCCCTTCAGGACATATTTTGGCATGCTGTTCACAGATGT 338
DB	184	GTCCACTGCTGCTGTTTCTCCCTTCAGGACATATTTTGGCATGCTGTTCACAGATGT 243
QY	339	ACCAATGCTCATGGAATACGAAATGGAACAGATGCTGGCAGTAGTGAACAGCTAGT 398
DB	244	ACCAATGCTCATGGAATACGAAATGGAACAGATGCTGGCAGTAGTGAACAGCTAGT 303
QY	399	GGCAGCCCTGTGAGGGTGTGCAATTTTCCCGACATCCAGGTGTTTGGCATCAGGGCA 458
DB	304	GGCAGCCCTGTGAGGGTGTGCAATTTTCCCGACATCCAGGTGTTTGGCATCAGGGCA 363
QY	459	GCTGATGGAATCTGTGTTTGTGGATGACAGTCATACAAATTATATAGATGTGTAGT 518
DB	364	GCTGATGGAATCTGTGTTTGTGGATGACAGTCATACAAATTATATAGATGTGTAGT 423
QY	519	GTTAAGATGGCTCTTGCGGCGCATATTTCTCTATATGAAAGCTTCTTGTCACT 578
DB	424	GTTAAGATGGCTCTTGCGGCGCATATTTCTCTATATGAAAGCTTCTTGTCACT 483
QY	579	GGCTTCCTCATGTGTGATTTAAACATGTGGATGATTAAGATGAGTGTCTGCATAGTAA 638
DB	484	GGCTTCCTCATGTGTGATTTAAACATGTGGATGATTAAGATGAGTGTCTGCATAGTAA 543
QY	639	AAAGCACAATGATCTTGGAAATTAAGTCTGCTGCGATTTTCTTCACAGCCAGTTTCTGATGGA 698
DB	544	AAAGCACAATGATCTTGGAAATTAAGTCTGCTGCGATTTTCTTCACAGCCAGTTTCTGATGGA 603
QY	699	GAACAAGGCTTCAAGTTTTTTCGATGGAATGATGATGTCAGAGATTTGGCAATGCAAAATT 758
DB	604	GAACAAGGCTTCAAGTTTTTTCGATGGAATGATGATGTCAGAGATTTGGCAATGCAAAATT 663
QY	759	TGATTTGTTTCTTTTACCATATCTT----- 784
DB	664	TGATTTGTTTCTTTTACCATATCTT----- 723
QY	785	----- 784
DB	724	GGGACATGTGCTCTGTTCTGGCTTGTGCTTTTCCATGATGGGACAGTACTATCTCA 783

QY	785	-----	785	-----	784
Db	784	GGGTCAGTGGATAAAGTCTGTCATAGTATATGATACATACTAATCTAGATAATATATCTTCACACA	843		
QY	785	-----	784	-----	
Db	844	TTGACTCAGCACACACAGGTATGTCAACACTGTGTCTTTTGACACCTAATACCTTTTACTT	903		
QY	785	-----	784	-----	
Db	904	GCTACTGGTTCATGACAAAAACAGTGAAACATCTGGCAATTTGACCTGGAAAACACTTTTGC	963		
QY	785	--AGCAAGGGCGACAGAACATCACGTGAAGCAATTTACCGAAGATTTGGTCAGAGAGGTC	842		
Db	964	CAGCAAGGGCGCACAGAACATCACGTGAAGCAATTTACCGAAGATTTGGTCAGAGAGAGAT	1023		
QY	843	GTCCTCAACATGGCTTTGTGCACAAGATTTAAAGATCTTTGTGTATTTTCAGATGAT	902		
Db	1024	GTCTCAACATGGCTTTGTGCACAAGATTTAAAGATCTTTGTGTATTTTCAGATGAT	1083		
QY	903	AACATTGATGGAAAAGAACTGTGTAATTTACAAAAGAAAGTCGTGGCTGATGATTTGAAA	962		
Db	1084	AACATTGATGGAAAAGAACTGTGTAATTTACAAAAGAAAGTCGTGGCTGATGATTTGAAA	1143		
QY	963	ATTGAATCTTACGACCTGGTAAAGTGTGTGAGAAAATTTGAAGACCTCAGAGCCAAAG	1022		
Db	1144	ATTGAATCTTACGACCTGGTAAAGTGTGTGAGAAAATTTGAAGACCTCAGAGCCAAAG	1203		
QY	1023	GTTAAATCCCTTTCTTCAGAAATTCCTGATGAATTTATATGTCCATTAACCTAGAAACTT	1082		
Db	1204	GTTAAATCCCTTTCTTCAGAAATTCCTGATGAATTTATATGTCCATTAACCTAGAAACTT	1263		
QY	1083	ATGAAAGATCCGGTTCATGCGATCGAGTGGCTATTCATTTGAAAAGGAAGCAATGGAAAT	1142		
Db	1264	ATGAAAGATCCGGTTCATGCGATCGAGTGGCTATTCATTTGAAAAGGAAGCAATGGAAAT	1323		
QY	1143	TGGATTCAGCAAAAAGAAACGTACAAAGTCCCATGACAAATCTTGTCTCTTCACGGGTA	1202		
Db	1324	TGGATTCAGCAAAAAGAAACGTACAAAGTCCCATGACAAATCTTGTCTCTTCACGGGTA	1383		
QY	1203	CTTACACCAATATAGACTCTGTGAAAATGCGCATCAATAGATGGCTGGAGACACACCAAAAG	1262		
Db	1384	CTTACACCAATATAGACTCTGTGAAAATGCGCATCAATAGATGGCTGGAGACACACCAAAAG	1443		
QY	1263	TAAA 1266			
Db	1444	TAAA 1447			
RESULT 8					
LOCUS	CQ497179	1996 bp	DNA	linear	PAT 30-JAN-2004
DEFINITION	Sequence 29046 from Patent WO0160860.				
ACCESSION	CQ497179				
VERSION	CQ497179.1	GI:41462815			
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1	Schlegel, R., Endege, W O. and Monahan, J E.			
AUTHORS		Genes differentially expressed in human prostate cancer and their			
TITLE		use			
JOURNAL		Patent: WO 0160860-A 29046 23-AUG-2001;			
FEATURES		Millennium Predictive Medicine, Inc. (US)			
source		1. 1996			
		location/Qualifiers			
		/organism="Homo sapiens"			
		/mol_type="unassigned DNA"			
		/db_xref="taxon:9606"			
ORIGIN					

Query Match	69.1%	Score	878.8	DB	6	Length	1996
Best Local Similarity	80.7%	Pred.	NC.48-227				
Matches	1166	Conservative	0	Mismatches	2	Indels	276
						Gaps	1
QY	99	CAAAATAAAGAACATGCTGAAACCTGATTAACAATTAGTGCATGATCATGCTGACGATGTCAC	158				
Db	4	CAAAATAAAGAACATGCTGAAACCTGATTAACAATTAGTGCATGATCATGCTGACGATGTCAC	63				
QY	159	TGCTGTGCTTCTCTCTTTTCCCTCTTGGCTACTTGCTCTTGGACAAACAAATTGGCTTG	218				
Db	64	TGCTGTGCTTCTCTCTTTTCCCTCTTGGCTACTTGCTCTTGGACAAACAAATTGGCTTG	123				
QY	219	TACTGTTACGAGCACTTACTGAACTGACATGTCACATTCCTCAATGAAGTTATACCTATGCT	278				
Db	124	TACTGTTACGAGCACTTACTGAACTGACATGTCACATTCCTCAATGAAGTTATACCTATGCT	183				
QY	279	GTCCACTGCTGCTGTTTCTCCCTTCAGGACATATTTTGGCATCGTTCAACAGATGCT	338				
Db	184	GTCCACTGCTGCTGTTTCTCCCTTCAGGACATATTTTGGCATCGTTCAACAGATGCT	243				
QY	339	ACCACTGCTCTATGGAATATCTGAATAATGACAGATGCTGGCAGTATGGAACAGCTTAGT	398				
Db	244	ACCACTGCTCTATGGAATATCTGAATAATGACAGATGCTGGCAGTATGGAACAGCTTAGT	303				
QY	399	GGCAGCCCTGTGAGGGTTTGGCAGTTTCCCAAGTCCACGTTGTTGGCATCAGGGGCA	458				
Db	304	GGCAGCCCTGTGAGGGTTTGGCAGTTTCCCAAGTCCACGTTGTTGGCATCAGGGGCA	363				
QY	459	GCTATGGAACCTGTGTTTGTGGAATGACAGTCATACAAATTATATAGATGTGTAGT	518				
Db	364	GCTATGGAACCTGTGTTTGTGGAATGACAGTCATACAAATTATATAGATGTGTAGT	423				
QY	519	GTTAAAGATGCTCCTTGGCGGATGTGATTTTCTCTAATGGAAGCTTCTTGTGCACT	578				
Db	424	GTTAAAGATGCTCCTTGGCGGATGTGATTTTCTCTAATGGAAGCTTCTTGTGCACT	483				
QY	579	GGCTCTCATGTGGATTTTAAACAGTGTGGATGATTAATAAGATGTGTGATAGTGA	638				
Db	484	GGCTCTCATGTGGATTTTAAACAGTGTGGATGATTAATAAGATGTGTGATAGTGA	543				
QY	639	AAAGCACATGATCTTGGAAATACCGTGGCGATTTTCTTCAACAGCAGTTCTGATGGA	698				
Db	544	AAAGCACATGATCTTGGAAATACCGTGGCGATTTTCTTCAACAGCAGTTCTGATGGA	603				
QY	699	GAACAGGCTCTTCAGTTTCTTTCGATCTGATCATGTGTGAGGATTTGCCAAGTCAAAAT	758				
Db	604	GAACAGGCTCTTCAGTTTCTTTCGATCTGATCATGTGTGAGGATTTGCCAAGTCAAAAT	663				
QY	759	TGATTTGTTTCTTTTAAACCATATCTT-----	784				
Db	664	TGATTTGTTTCTTTTAAACCATATCTTAAAGTTTGAATTAATAATTAAGTACACTGAGT	723				
QY	785	-----	784				
Db	724	GGGACATGTGCTCCTGTTCTGGCTTGCTTTTCCCATGATGGCGAGTCTAGTCTGA	783				
QY	785	-----	784				
Db	784	GGGTCAAGTGAATAGTCTGTGATATAGTATGATTAATACAGAAATATACTTCACACA	843				
QY	785	-----	784				
Db	844	TTGATCAGCACACCGAGTATGTCAACAATTGTGCTTTTGACACTAATACCTTTTACTT	903				
QY	785	-----	784				
Db	904	GCTACTGATTCAATGACAAAACAGTGAACATCTGGCAATTTGACCTGGAAAACATTGTC	963				
QY	785	--AGCAAGGCGACAGAACATCAGCTGAAGCAATTTACGAGATTTGTCAGAGAGGTC	842				
Db	964	CAAGMAAGGCGACAGAACATCAGCTGAAGCAATTTACGAGATTTGTCAGAGAGGAT	102				
QY	843	GTTCTCAACATGCTTTGTGACACAAGATTTAAAGATCTTGTGTATTTTCAAGATGAAT	902				

Db	1024	GTCTCAACATGCGCTTTGTGCAACAGATTTAAGAATCTTGTTGGTATTTTCAAGATGAAT	10833
Oy	903	AACATTGATGAAAAAGAACTGTTGAATCTTACAAAAGAAAGCTGCGCTGATGATTTGAAA	962
Db	1084	AACATTGATGAAAAAGAACTGTTGAATCTTACAAAAGAAAGCTGCGCTGATGATTTGAAA	11434
Oy	963	ATTGAATCTCTAGGACCTGCGTGAATGTAAGAGTGCTGAGAAAAATTGAAGAGCTTACAGACCAG	10222
Db	1144	ATTGAATCTCTAGGACCTGCGTGAATGTAAGAGTGCTGAGAAAAATTGAAGAGCTTACAGACCAG	12033
Oy	1023	GTTAAATCCCTTCTTCAGGAATCCGATGAATTTATATGTCCAACTACTAGAGAACTT	10822
Db	1204	GTTAAATCCCTTCTTCAGGAATCCGATGAATTTATATGTCCAACTACTAGAGAACTT	12633
Oy	1083	ATGAAGAATCCGCTGATCGCATCAGATGCGCTATTCAATGAAAAAGAAAGCAATGGAANAAT	11422
Db	1264	ATGAAGAATCCGCTGATCGCATCAGATGCGCTATTCAATGAAAAAGAAAGCAATGGAANAAT	13222
Oy	1143	TGATTCAGCAAAAAAGAAACGTACAAGTCCCATGACAAATCTTGTTCTTCTTCACGCGTA	12022
Db	1324	TGATTCAGCAAAAAAGAAACGTACAAGTCCCATGACAAATCTTGTTCTTCTTCACGCGTA	13833
Oy	1203	CTTACACCAATATAGGACTCTGAAAAATGGCCATCATATGATGGCTGAGACACACCAACAAAG	12622
Db	1384	CTTACACCAATATAGGACTCTGAAAAATGGCCATCATATGATGGCTGAGAGACACACCAACAAAG	14433
Oy	1263	TAAA 1266	
Db	1444	TAAA 1447	

AX077675	RESULT 9
LOCUS	AX077675 1291 bp DNA linear PAT 22-FEB-2001
DEFINITION	Sequence 130 from Patent WO0105970.
ACCESSION	AX077675
VERSION	AX077675.1 GI:13122057
KEYWORDS	
SOURCE	
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE	1 Yue, H., Tang, Y. T., Bandman, O., Hillman, J. L., Lal, P., Au-Young, J., Reddy, R., Yang, J., Baughn, M. R., Lu, D. A., Azimzai, Y. and Paterson, C.
JOURNAL	Gp-binding protein associated factors
FEATURES	Patent: WO 0105970-A 10 25-JAN-2001;
SOURCE	Incyte Genomics, Inc. (us)
	Location/Qualifiers
	1..1291
	/organism="Homo sapiens"
	/mol_type="unassigned DNA"
	/db_xref="taxon:9606"
	/note="Incyte ID No: 4546403CB1"
ORIGIN	
Query Match	37.6%; Score 478.6; DB 6; Length 1291;
Best Local Similarity	98.2%; Pred. No. 1.6e-118;
Match 484; Conservative	0; Mismatches 9; Indels 0; Gaps 0;
QY	774 ACCCATATCTTAGCAAGGCGACAGACATAGCTGGAAGCAATTACCGAAGATTGGTCA 833
DB	547 ACACTTGGCCAGCAAGAGACACAGAACATAGCTGGAAGCAATTACCGAAGATTGGTCA 606
QY	834 GAGGAGGTGCTCTCAACATGGCTTTGGTGCACAAAGATTAAAGATCTTGGTGGTATTTTC 893
DB	607 GAGGAGGATGTCCTCAACATGGCTTTGGTGCACAAAGATTAAAGATCTTGGTGGTATTTTC 666
QY	894 AAGATGAATACATTGATGTAAGAAAGAACTGTGTAATCTTACAAAGAAAGCTGGCTGAT 953
DB	667 AAGATGAATACATTGATGTAAGAAAGAACTGTGTAATCTTACAAAGAAAGCTGGCTGAT 726

QY 954 GATTGAAATGATCTCTAGAGCTCGCTAGTAAAGTCTGAGAAATGGAAGCTC 1013
 DB 727 GATTGAAATGATCTCTAGAGCTCGCTAGTAAAGTCTGAGAAATGGAAGCTC 786
 QY 1014 AGGACCAAGTTAAATCCCTTTCTTCAGAAATTCCTGATGAATTTATATGCTCAATPACT 1073
 DB 787 AGACCAAGTTAAATCCCTTTCTTCAGAAATTCCTGATGAATTTATATGCTCAATPACT 846
 QY 1074 AGGAACTTATGAAAGTCCGCTATCGATCGATGAGTCTTATCATATGAAAGGAAGCA 1133
 DB 847 AGGAACTTATGAAAGTCCGCTATCGATCGATGAGTCTTATCATATGAAAGGAAGCA 906
 QY 1134 ATGGAATTTGATTCAGCAAAAAGAAACGTACAAAGTCCCATGACAAATCTTGTCTTCTCT 1193
 DB 907 ATGGAATTTGATTCAGCAAAAAGAAACGTACAAAGTCCCATGACAAATCTTGTCTTCTCT 966
 QY 1194 TCAGCGCTACTTACACCAATAGGACTCTGAAAATGGCCATCAATGATGCTGAGACCA 1253
 DB 967 TCAGCGCTACTTACACCAATAGGACTCTGAAAATGGCCATCAATGATGCTGAGACCA 1026
 QY 1254 CACCAAAAGTAAA 1266
 DB 1027 CACCAAAAGTAAA 1039

RESULT 10
 LOCUS BC061948 2130 bp mRNA linear ROD 19-NOV-2003
 DEFINITION Mus musculus RIKEN cDNA 2610014F08 gene, mRNA (cDNA clone MGC:65547
 IMAGE:6312489), complete cds.
 ACCESSION BC061948
 VERSION BC061948.1 GI:38303926
 KEYWORDS MGC.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 2130)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.L., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stappleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Uedlin, T.B., Tomshyuk, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullan, S.J., Bosak, S.A., McGowan, P.U., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hui, S.W., Vollbrecht, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettner, M., Madan, A., Young, A.C., Shevchenko, Y., Sanchez, A., Whiting, M., Madan, A., Touchman, J.W., Green, E.D., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, J., D.B., Schermer, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

TITLE
 JOURNAL MEDLINE
 PUBMED 22388257
 REFERENCE 2 (bases 1 to 2130)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (10-NOV-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
 COMMENT NHA-MGC Project URL: <http://mgc.ncl.nih.gov>
 Contact: MGC help desk
 Email: cgaps-r@mail.nih.gov
 Tissue Procurement: Susan L. Sullivan, PhD.

CDNA Library Preparation: ResGen, Invitrogen Corp
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisec.nih.gov/>
 Contact: nisc_mgc@ncl.nih.gov
 Akher, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Grant, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R., Maduro, Q.L., Mastaglio, C., Misker, B., Mestrian, S.D., McLooney, J.C., McDowell, J., Pearson, R., Stantrop, S., Thomas, P.J., Touchman, J.W., Tsurgenev, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAP Plate: 127 Row: 1 Column: 1
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES

source location/Qualifiers
 1..2130
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="MGC:65547 IMAGE:6312489"
 /rname_type="Olfactory epithelium, neonatal mouse, CS7BL/6"
 /clone_11b="NIH MGC_129"
 /lab_host="DH10B"
 /note="Vector: PCMV-SPORT6"
 1..2130
 /gene="2610014F08R1k"
 /note="Bynonym: 1700048B19R1k"
 /db_xref="LocustID:72137"
 /db_xref="MGI:1919387"
 208..1632
 /gene="2610014F08R1k"
 /translation="MVRLLITLADGDGVSCCAFAAALATSLDKTRIVLSPPAE LVSPLKRTTIVVHCCSPSPSGVTLASGSTDYTLVMSHSHRTITVLEQPGSPYRV CCFSPSAYLASGADGSLALNNAQTYKYRCGSVDSLVACAFSPDGLVTSISG GUTVWDMDRMLHSEKADLIGITCSFSSQPSGEGSLQSLACGQDCBIKLMAV TITRVVGFELKXKSTLSGHCAPVLACAFSDHDKMLASGSDSVIIGHIGPOSVHTL TQTRVTVTCAPAPNTLLATSGMKTVMIQFDLETPCGSMNPDLKHITRESEED VSVVRAQGLIEDLVNGIFRANNTDGEKLIHLITRESIAGDKITSLGSRVLSIEEL AKMDLSSSGLIPDFICPTIRELMKQPVLASDGYSTREAMESHIKKRTSPMTIAL PSLVLPNTLTLMAINRWLETHK"
 217..1131
 /gene="2610014F08R1k"
 /note="WP40: Region: WD40 domain, found in a number of eukaryotic proteins that cover a wide variety of functions including adaptor/regulatory modules in signal transduction, pre-mRNA processing and cytoskeleton assembly"
 /db_xref="CD:cd00200"
 1192..1383
 /gene="2610014F08R1k"
 /note="SAM: Region: Sterile alpha motif."
 /db_xref="CD:cd00166"
 1408..1629
 /gene="2610014F08R1k"
 /note="U-box: Region: U-box domain. This domain is related to the Ring finger pfam0097 but lacks the zinc binding residues"
 /db_xref="CD:pfam04564"

ORIGIN

Query Match 37.3%; Score 474.6; DB 10; Length 2130;
 Best Local Similarity 62.8%; Pred. No. 2.1e-117;
 Matches 955; Conservative 0; Mismatches 289; Indels 277; Gaps 3;

19 CGCGGCAAGTACCGGCAACCGCCGCGGCACTTGAAGCGGATCCCGCGCGCCGC 78
 117 CTGACACCCCGGAGCGCGGAGGAGGAGCGCCCGGAGGCGGAGCTCTAGTCCGGTG 176
 79 TCTGCAAGGCTGTTTTCTTCAATAAAGAACATGCTGAACCTGATTCACATTAAGTG 138
 177 CGAGGGAAGCTG-TTACTTTGTGTGAGGAACATGCTGAGCTGATTCACACGCTGGCTG 235
 139 ATCATGAGGAGATGTCATGCTGTGCTTCTCTTTCCCTTGTGGCTACTGCTCT 198
 236 ATCAAGGATGATGCTGAGCTGCTGCGCTTCTGGCGCTCTCTGCGCACCTGCTCT 295
 199 TGAACAAACAAATTCGCTGATCTGCTGATGCTGATGCTGATGCTGATGCTGAT 258
 296 TGAACAAACAAATTCGCTGATCTGCTGATGCTGATGCTGATGCTGATGCTGAT 355
 259 TGAAGTTTCAATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 318
 356 TGAAGTTTCAATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 415
 319 CATGCTGTTCAACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 378
 416 CATGCTGTTCAACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 475
 379 CAGTATGAGAACAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 438
 476 CCGTGTGAGAACAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 535
 439 CGTGTGTCATCAGGCGCAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 498
 536 CCACTAGCTGCTGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 595
 499 AATATATAGATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 558
 596 AACTATATAGATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 655
 559 ATGGAAGCTTCTTGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 618
 656 ATGGAAGCTTCTTGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 715
 619 TGAAGTGTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 678
 716 TGAAGTGTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 775
 679 CACAGCCAGTTTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 738
 776 CACAGCCCTCTCTGCGCGAG---AAGGCTTCCAGTCTTACAGTGTGCGTCACTGTGCTC 832
 739 AGATGTCAGTCAATAATTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 786
 833 AAGACTGTGAATCAAACTCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 892
 787 ----- 786
 893 AATATATAAGACACTAAGTGGGACCTGCGCCCTGTTCTGCGCTGTGCTTTTTCACATG 952
 787 ----- 786
 953 ATGAAAGATCTTGCATCGGGGCTCAGTGAATAATCTGTGATCATACATGATTCGGCC 1012
 787 ----- 786
 1013 CTCAGAGTGTCTACACGCTGACTGACATACAGAGTATGTAACACTGTGTGCGTTTG 1072
 787 ----- 786
 1073 CACCAACACTCTCTTACTGCTACTGCTCATGATGACAGACAGTGAACATTTGGCAGT 1132

787 -----CAAGGCGCAGACAGAACTACAGCTGACGCAATTTACCGAG 825
 1133 TTGACTGGAACACCTTGGCCAGAGACATGAACGACCCGCTGAACATTTCTGATGAG 1192
 826 ATTGTCAGAGAGAGTGTCTCAACATGCTTTGTGCAACAATTTAAAGATCTTGTG 885
 1193 AATGTCAGAGAGAGATGTCTCCGTGTGTGTGTGCTCAAGGCTTGAAGACTCGTGC 1252
 886 GTATTTTCAAGATGAATTAACATGATGATGATGATGATGATGATGATGATGATGATG 945
 1253 GTATTTTCAAGGCGCAGACATGATGATGATGATGATGATGATGATGATGATGATGATG 1312
 946 TGGCTGATGATTTGAAATTTGAATCTGATGATGATGATGATGATGATGATGATGATGATG 1005
 1313 TGGCTGATGATTTGAAATTTGAATCTGATGATGATGATGATGATGATGATGATGATGATG 1372
 1006 AAGAGCTCAGAGACCAAGTTAAATCCCTTCTTCAAGAAATTCATGATGAATTTATGTC 1065
 1373 AAGAGCTCAGAGACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1432
 1066 CAATTAACAGAGACCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1125
 1433 CAATTAACAGAGACCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1492
 1126 AAGAGCAATGAGAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1185
 1493 GAGAGCAATGAGAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1552
 1186 TTCTTCTTCAAGGCGCTTCTTCAACCAATTAAGATGATGATGATGATGATGATGATGAT 1245
 1553 CTCTCCCTTCACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1612
 1246 TGAAGACACACCAAAAGTAA 1266
 1613 TGAAGACACACCAAAAGTAA 1633

RESULT 11
 BC050792 1547 bp mRNA linear ROD 06-MAY-2003
 LOCUS Mus musculus RIKEN CDNA 2610014F08 gene, mRNA (cDNA clone
 IMAGE:6742025), partial cds.
 ACCESSION BC050792
 VERSION BC050792.1 GI:30048136
 KEYWORDS Mus musculus (house mouse)
 SOURCE Mus musculus
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;
 1 (bases 1 to 1547)
 REFERENCE Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L.H., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
 Diatchenko, L., Marisina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butterfield, Y.S., Krzywicki, M.I., Skalski, U., Smailus, D.E.,
 Scherch, A., Schein, J.E., Jones, S.G., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 MEDLINE 22388257

FEATURES	
source	Location/Qualifiers
	1. .147021
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/clone="MRP.35_A1"
	/clone_lib="RPCI-11 human BAC library"
repeat_region	complement(382. .683)
repeat_region	/rpt_family="AluSx"
repeat_region	1233. .1446
repeat_region	/rpt_family="L1MB3A"
repeat_region	4058. .4138
repeat_region	/rpt_family="MER96"
repeat_region	complement(4768. .5069)
repeat_region	/rpt_family="AluY"
unsure	5268. .5687
	/note="Single-stranded coverage."
repeat_region	5392. .5693
repeat_region	/rpt_family="AluSg"
repeat_region	complement(5856. .6314)
repeat_region	/rpt_family="L1MC5"
repeat_region	7493. .7713
repeat_region	/rpt_family="MIR"
repeat_region	complement(9050. .9111)
repeat_region	/rpt_family="L2"
repeat_region	complement(9768. .9867)
repeat_region	/rpt_family="AluJo/FRAM"
repeat_region	9927. .10125
repeat_region	/rpt_family="AluSg/X"
repeat_region	complement(11036. .11131)
repeat_region	/rpt_family="MIR"
repeat_region	12467. .12556
repeat_region	/rpt_family="MIR"
repeat_region	complement(13039. .13158)
repeat_region	/rpt_family="MIR"
repeat_region	13216. .13368
repeat_region	/rpt_family="MERSB"
repeat_region	complement(13631. .13682)
repeat_region	/rpt_family="MIR"
repeat_region	complement(14013. .14344)
repeat_region	/rpt_family="AluSg"
repeat_region	complement(14881. .15068)
repeat_region	/rpt_family="L2"
repeat_region	complement(15777. .15920)
repeat_region	/rpt_family="FLAM_A"
repeat_region	15977. .16097
repeat_region	/rpt_family="MERSA"
repeat_region	complement(16111. .16422)
repeat_region	/rpt_family="AluSx"
repeat_region	complement(16435. .16990)
repeat_region	/rpt_family="L1MB2"
repeat_region	complement(16991. .17293)
repeat_region	/rpt_family="AluSc"
repeat_region	complement(17318. .17483)
repeat_region	/rpt_family="FAM"
repeat_region	complement(17491. .17574)
repeat_region	/rpt_family="L1MB2"
repeat_region	complement(17578. .17632)
repeat_region	/rpt_family="MERSB"
repeat_region	18398. .18544
repeat_region	/rpt_family="MIR"
repeat_region	complement(18601. .18892)
repeat_region	/rpt_family="AluSx"
repeat_region	21078. .21102
repeat_region	/rpt_family="(CA)n"
repeat_region	complement(21331. .21814)
repeat_region	/rpt_family="MLTID"
repeat_region	complement(24151. .24346)
repeat_region	/rpt_family="MLTIB1"
repeat_region	complement(26673. .26977)
repeat_region	/rpt_family="AluSc"
repeat_region	complement(27038. .27349)
repeat_region	/rpt_family="AluB"

repeat_region	27424. .27614
repeat_region	/rpt_family="MERS3"
repeat_region	28893. .29218
repeat_region	/rpt_family="MER7A"
repeat_region	30893. .31200
repeat_region	/rpt_family="AluSg"
repeat_region	31226. .31363
repeat_region	/rpt_family="MIR"
repeat_region	31429. .31608
repeat_region	/rpt_family="MIR"
repeat_region	31694. .31715
repeat_region	/rpt_family="AT rich"
repeat_region	complement(32264. .32559)
repeat_region	/rpt_family="MER7A"
repeat_region	34027. .34056
repeat_region	/rpt_family="(TTTG)n"
repeat_region	complement(34057. .34340)
repeat_region	/rpt_family="AluSc"
repeat_region	complement(34350. .34643)
repeat_region	/rpt_family="AluSx"
repeat_region	complement(35919. .36110)
repeat_region	/rpt_family="MER34B"
repeat_region	36113. .36414
repeat_region	/rpt_family="AluSg"
repeat_region	complement(36626. .36713)
repeat_region	/rpt_family="MER34B"
repeat_region	37317. .37614
repeat_region	/rpt_family="AluSx"
repeat_region	complement(39201. .39613)
repeat_region	/rpt_family="L1MB3"
repeat_region	complement(40549. .40933)
repeat_region	/rpt_family="MLTIG1"
repeat_region	41531. .41566
repeat_region	/rpt_family="(TTTG)n"
repeat_region	43451. .43494
repeat_region	/rpt_family="(TG)n"
repeat_region	complement(43767. .44074)
repeat_region	/rpt_family="AluSg"
repeat_region	complement(44084. .44352)
repeat_region	/rpt_family="AluB"
repeat_region	complement(44359. .44541)
repeat_region	/rpt_family="L1MA6"
repeat_region	44542. .44872
repeat_region	/rpt_family="AluSg"
repeat_region	complement(44873. .45183)
repeat_region	/rpt_family="L1MA6"
repeat_region	complement(45708. .46376)
repeat_region	/rpt_family="L1MB2"
repeat_region	complement(46377. .46677)
repeat_region	/rpt_family="AluSx"
repeat_region	complement(46678. .46940)
repeat_region	/rpt_family="L1MB2"
repeat_region	complement(48115. .48193)
repeat_region	/rpt_family="MERSB"
repeat_region	complement(48485. .48681)
repeat_region	/rpt_family="MIR"
repeat_region	49373. .49660
repeat_region	/rpt_family="AluSx"
repeat_region	complement(50181. .50391)
repeat_region	/rpt_family="MERS8A"
repeat_region	52831. .52908
repeat_region	/rpt_family="MIR"
repeat_region	57191. .57478
repeat_region	/rpt_family="AluSc"
repeat_region	57516. .57795
repeat_region	/rpt_family="AluYB8"
repeat_region	57983. .58280
repeat_region	/rpt_family="AluSg"
repeat_region	60124. .60147
repeat_region	/rpt_family="(TTTG)n"
repeat_region	complement(60148. .60393)
repeat_region	/rpt_family="AluSg"

Query Match 33.8%; Score 430; DB 9; Length 147021;
 Best Local Similarity 96.7%; Pred. No. 4,18-105;
 Matches 439; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 82 TGCAGGCTGTTTCTTCAATAAAGAACATGGTGAACATGATTCACATTTAGCTATC 141
 DB 106602 TTCAGGCTGTTTCTTCAATAAAGAACATGGTGAACATGATTCACATTTAGCTATC 106543

QY 142 ATGGTGAAGATGCAATGCTGCTGCTCTTCCCTTCCCTTGGCTACTTGCCTCTGG 201
 DB 106542 ATGGTGAAGATGCAATGCTGCTGCTCTTCCCTTGGCTACTTGCCTCTGG 106483

QY 202 ACAAAACAAATTCGCTGATCTGTTAGCTGATCTTACGAACTCCCATTCCTCATGA 261
 DB 106482 ACAAAACAAATTCGCTGATCTGTTAGCTGATCTTACGAACTCCCATTCCTCATGA 106423

QY 262 AGTTTCATACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 321
 DB 106422 AGTTTCATACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 106363

QY 322 CGTGTTCACAGATGATGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 381
 DB 106362 CGTGTTCACAGATGATGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 106303

QY 382 TGAATGAACAGCCTAGTGGCAGCCCTGGAAGGTTTGCAGTTTCCCGAGACTCCAGT 441
 DB 106302 TGAATGAACAGCCTAGTGGCAGCCCTGGAAGGTTTGCAGTTTCCCGAGACTCCAGT 106243

QY 442 GTTTCGATCAGGGGCGAGCTGATGAACTGTGTTTGTGGAATGCAAGTATCAAAAT 501
 DB 106242 GTTTCGATCAGGGGCGAGCTGATGAACTGTGTTTGTGGAATGCAAGTATCAAAAT 106183

QY 502 TATATAGATGCTGATGCTTAAAGATGCTCTCT 535
 DB 106182 TATATAGATGCTGATGCTTAAAGATGCTCTCTCT 106149

RESULT 14
 AC008277/c 177089 bp DNA linear PRI 29-OCT-2002
 LOCUS Homo sapiens BAC clone RP11-311B14 from 2, complete sequence.
 DEFINITION AC008277
 AC008277
 VERSION AC008277.4 GI:10716648
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 177089)
 Toward a complete human genome sequence
 Sulston, J.E. and Waterston, R.
 JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
 MEDLINE 9847074
 PUBMED 99063792
 REFERENCE 2 (bases 1 to 177089)
 AUTHORS Kyung, K., Maupin, R. and Hawkins, M.
 TITLE The sequence of Homo sapiens BAC clone RP11-311B14
 JOURNAL Unpublished (2001)
 REFERENCE 3 (bases 1 to 177089)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (31-JUL-1999) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 REFERENCE 4 (bases 1 to 177089)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (27-APR-2000) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 REFERENCE 5 (bases 1 to 177089)
 AUTHORS Waterston, R.
 TITLE Direct Submission

JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Submitted (07-OCT-2000) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 6 (bases 1 to 177089)
 Waterston, R.H.
 Direct Submission
 Submitted (25-MAR-2001) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 7 (bases 1 to 177089)
 Waterston, R.
 Direct Submission
 Submitted (29-OCT-2002) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Oct 7, 2000 this sequence version replaced gi:7656634.

Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc
 Contact: sapiens@watson.wustl.edu
 ----- Summary Statistics
 Center project name: H_NH0311B14

NOTICE: This sequence may not represent the entire insert of this
 clone. It may be shorter because we only sequence overlapping
 clone sections once, or longer because we provide a small overlap
 between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate
 chemistry, or covered by high quality data (i.e., phred quality >=
 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by sequence
 from more than one subclone; and the assembly was confirmed by
 restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D.
 McPherson, Department of Genetics, Washington University, St. Louis
 MO. For additional information about the map position of this
 sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RP11-11 human BAC library was made from the blood of one male
 donor, as described by Osseogawa, K., Moon, P.Y., Zhao, B., Frangen, E.,
 Tateo, M., Catane, J.J., and de Jong, P.J. (1998) An improved
 approach for construction of bacterial artificial chromosome
 libraries. *Genomics* 51:1-8. The clone may be obtained either from
 Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong
 and coworkers at <http://www.chori.org>
 VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the right is RP11-542H1, 200 base pair
 overlap. The clone sequenced to the left is RP11-292A10, 200 base
 pair overlap. Actual start of this clone is at base position 1 of
 RP11-311B14; actual end is at base position 23376 of RP11-542H1.
 RP11-311B14 contains a transposon in the growth of the clone which
 has not been included in the submitted sequence. The transposon
 would insert after base position 212.

RP11-311B14 contains sequence from base position 23113 to 25636
 which is not represented in the full sequence of its neighbor
 RP11-292A10.

FEATURES

source
 1..177089
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="2"
 /map="2"
 /clone="RP11-311B14"

```
/clone_lib="RPCI-11"
repeat_region 65..146
/rpc_family="L1"
repeat_region 185..596
/rpc_family="L1"
repeat_region 728..945
/rpc_family="L1"
repeat_region 959..1171
/rpc_family="L1"
repeat_region 1365..1969
/rpc_family="L1"
misc_feature 2185..2293
/note="similar to EST AL040518 (NID:g5409467)"
repeat_region 2484..3017
/rpc_family="ERVX"
repeat_region 3658..3694
/rpc_family="L2"
repeat_region 3915..4216
/rpc_family="Alu"
repeat_region 4720..4861
/rpc_family="Alu"
repeat_region 5397..5708
/rpc_family="L1"
misc_feature 6044..6136
/note="similar to EST AL040518 (NID:g5409467)"
repeat_region 6554..6851
/rpc_family="Alu"
repeat_region 7237..7513
/rpc_family="Alu"
repeat_region 8895..9148
/rpc_family="L2"
repeat_region 9369..9554
/rpc_family="MERI_type"
misc_feature 10258..10443
/note="similar to EST AA118718 (NID:g1676210) mp57c07.r1"
misc_feature 10259..10452
/note="similar to EST AL040518 (NID:g5409467)"
misc_feature 10353..10443
/note="similar to EST AA1189142 (NID:g3740351) qd04a04.x1"
repeat_region 10908..11207
/rpc_family="Alu"
repeat_region 11747..11904
/rpc_family="Alu"
repeat_region 12188..12490
/rpc_family="Alu"
repeat_region 12696..12988
/rpc_family="Alu"
misc_feature 13170..13582
/note="similar to EST AA1189142 (NID:g3740351) qd04a04.x1"
misc_feature 13170..13411
/note="similar to EST AA118718 (NID:g1676210) mp57c07.r1"
misc_feature 13370..13429
/note="similar to EST AA501581 (NID:g2236548) ne97g11.g1"
misc_feature 13284..13544
/note="similar to EST AA482240 (NID:g7052346)"
repeat_region 13117..13894
/rpc_family="MERI_type"
repeat_region 13913..14066
/rpc_family="Alu"
repeat_region 14178..14303
/rpc_family="MERI_type"
repeat_region 14308..14608
/rpc_family="Alu"
repeat_region 15230..15355
/rpc_family="L2"
repeat_region 15345..15679
/rpc_family="L2"
repeat_region 15924..16201
/rpc_family="L2"
repeat_region 16595..17554
/rpc_family="Alu"
misc_feature /note="CpG island (4GC=68.2, o/e=0.80, #CpGs=95)"
18311..18637
/note="similar to EST AW170043 (NID:g6401568) xj36b07.x1"
```

```
repeat_region 18432..18642
/rpc_family="L1"
repeat_region 18673..18731
/rpc_family="L1"
repeat_region 18732..19036
/rpc_family="Alu"
repeat_region 19037..19162
/rpc_family="L1"
repeat_region 19206..19463
/rpc_family="Alu"
repeat_region 19869..20164
/rpc_family="Alu"
repeat_region 20283..20576
/rpc_family="Alu"
repeat_region 20614..20759
/rpc_family="L1"
repeat_region 21115..21187
/rpc_family="L1"
repeat_region 21800..22026
/rpc_family="Alu"
```

Query Match 33.8%; Score 430; DB 9; Length 177089;

Best Local Similarity 96.7%; Pred No. 4.2e-105; Indels 0; Gaps 0;

Matches 439; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 82 TGCAGGCTGTTTCTTCAATTAAGACATGATGAACTGATTCACACATTAGCTGATC 141

Db 13596 TTCAGGCTGTTTCTTCAATTAAGACATGATGAACTGATTCACACATTAGCTGATC 13537

QY 142 ATGCTACGATGTCACATGCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 201

Db 13536 ATGCTACGATGTCACATGCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 13477

QY 202 ACAAAACAATTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 261

Db 13476 ACAAAACAATTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 13417

QY 262 AGTTTCATCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 321

Db 13416 AGTTTCATCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 13357

QY 322 CGTGTTCACACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 381

Db 13356 CGTGTTCACACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 13297

QY 382 TGATGACACGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 441

Db 13296 TGATGACACGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 13237

QY 442 GTTTCGATCAGGCGCAGCTGATGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 501

Db 13236 GTTTCGATCAGGCGCAGCTGATGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 13177

QY 502 TATATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 535

Db 13176 TATATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 13143

RESULT 15

CQ473162/c 446 bp DNA linear PAT 30/JAN-2004

LOCUS DEFINITION Sequence 5029 from Patent WO0160860

ACCESSION CQ473162

VERSION CQ473162.1 GI:41438781

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Top 100

REFERENCE 1 Schlegel, R., Endege, W.O. and Monahan, J.E. Genes differentially expressed in human prostate cancer and their

JOURNAL Patent: WO 0160860-A 5029 23-AUG-2001;
 FEATURES Millennium Predictive Medicine, Inc. (US)
 Location/Qualifiers
 source 1..446
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 26.3%; Score 335; DB 6; Length 446;
 Best Local Similarity 96.3%; Pred. No. 1.4e-79;
 Matches 365; Conservative 0; Mismatches 10; Indels 4; Gaps 2;

QY	412	GGGTTTGCAGTTTCCCGACAGTCCAGTG--TTGGCATCAGGGGACGTGAGAA	468
DB	445	GGTTTGCAGTTTTCCTCCAGACTCCAGCTGTTTGGCATCAGGGGACGTGAGAA	386
QY	469	CTGTGTTTGTGGAATGCACAGTCATACAAATTATATAGATGTGTAGTTTAAAGATG	528
DB	385	CTGTGTTTGTGGAATGCACAGTCATACAAATTATATAGATGTGTAGTTTAAAGATG	326
QY	529	GCTCCTTGGCGGCAATGTCATTTTCTCTAATGGAAGCTTTTGTCACTGGCTCTCAT	588
DB	325	GCTCCTTGGCGGCAATGTCATTTTCTCTAATGGAAGCTTTTGTCACTGGCTCTCAT	266
QY	589	GTGATGTTTAAACAGTGTGGGATGATAAATGAGGTGTGTCATAGTG-AAAAGCAAT	647
DB	265	GTGATGTTTAAACAGTGTGGGATGATAAATGAGGTGTGTCATAGTGAAAAGCAAT	206
QY	648	GATCTTGAATTAACCTGCTGCGATTTTCTTCAAGCCAGTTTCTGATGAGAACAAAGT	707
DB	205	GATCTTGAATTAACCTGCTGCGATTTTCTTCAAGCCAGTTTCTGATGAGAACAAAGT	146
QY	708	CTTCAGTTTTCAGTGGCATCATGTGTGAGATTGCCAAGTCAAAATTGAGTTGTT	767
DB	145	CTTCAGTTTTCAGTGGCATCATGTGTGAGATTGCCAAGTCAAAATTGAGTTGTT	86
QY	768	TCCTTTACCCATATCTTAG	786
DB	85	TCCTTTACCCATATCTTAG	67

Search completed: February 5, 2005, 07:01:54
 Job time : 5676 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 5, 2005, 13:45:37 ; Search time 5684 Seconds
(without alignments)
3273.540 Million cell updates/sec

Title: US-10-077-111-13
Perfect score: 2047
Sequence: 1 MVKLHTLADHDGDDVNCAP.....LTPNRTKMAINRWLETHK 384

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 segs, 24227607955 residues

Total number of hits satisfying chosen parameters: 2805242

Minimum DB seq length: 0
Maximum DB seq length: 200

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlh
-Q=/cg2_1/USPTO.spool/US10077111/runat_04022005_121141_7181/app.query.fasta.1.583
-DB=GenEmbl -QFMT=faetap -SUFPTX=p2n2lm200.rge -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=bl0sum62 -TRANS=human40.cdl
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptlo -NOBM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200
-USER=US10077111.QCGN.1.14200 @runat_04022005_121141_7181 -MCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb_ba.*
2: gb_hg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
C 1	300	14.7	180	CQ110706
C 2	300	14.7	180	CQ149443
C 3	300	14.7	180	CQ232723
C 4	300	14.7	180	CQ270754

C 5	300	14.7	180	6	CQ308036	CQ308036 Sequence
C 6	300	14.7	180	6	CQ344856	CQ344856 Sequence
C 7	210	10.3	132	6	CQ058095	CQ058095 Sequence
C 8	210	10.3	132	6	CQ077416	CQ077416 Sequence
C 9	210	10.3	132	6	CQ108429	CQ108429 Sequence
C 10	210	10.3	132	6	CQ147070	CQ147070 Sequence
C 11	210	10.3	132	6	CQ206855	CQ206855 Sequence
C 12	210	10.3	132	6	CQ230294	CQ230294 Sequence
C 13	210	10.3	132	6	CQ268431	CQ268431 Sequence
C 14	210	10.3	132	6	CQ305465	CQ305465 Sequence
C 15	85	4.2	131	6	AX905781	AX905781 Sequence
C 16	85	4.2	131	6	BD041314	BD041314 Sequence
C 17	82	4.0	173	6	CQ114451	CQ114451 Sequence
C 18	82	4.0	173	6	CQ153330	CQ153330 Sequence
C 19	82	4.0	173	6	CQ236619	CQ236619 Sequence
C 20	82	4.0	173	6	CQ274216	CQ274216 Sequence
C 21	82	4.0	173	6	CQ311386	CQ311386 Sequence
C 22	82	4.0	173	6	CQ348524	CQ348524 Sequence
C 23	79.5	3.9	169	6	AR269854	AR269854 Sequence
C 24	77.5	3.8	163	11	G18760	G18760 BMS960 cow
C 25	73	3.6	190	6	AX099366	AX099366 Sequence
C 26	73	3.6	190	6	BD021982	BD021982 Sequence
C 27	66	3.2	104	9	HSEXI4B01	X92500 H. sapiens t
C 28	66	3.2	154	6	AR245693	AR245693 Sequence
C 29	66	3.2	168	8	AJ719154	AJ719154 Nicotiana
C 30	65	3.2	131	6	CQ080172	CQ080172 Sequence
C 31	65	3.2	131	6	CQ113458	CQ113458 Sequence
C 32	65	3.2	131	6	CQ152338	CQ152338 Sequence
C 33	65	3.2	131	6	CQ185706	CQ185706 Sequence
C 34	65	3.2	131	6	CQ235679	CQ235679 Sequence
C 35	65	3.2	131	6	CQ273247	CQ273247 Sequence
C 36	65	3.2	131	6	CQ310696	CQ310696 Sequence
C 37	65	3.2	131	6	CQ347535	CQ347535 Sequence
C 38	64	3.1	136	14	AY366579	AY366579 Hepatitis
C 39	63	3.1	102	9	S65996	S65996 NK-TR-putat
C 40	63	3.1	135	14	AY366585	AY366585 Hepatitis
C 41	62.5	3.1	135	9	AB0046832	AB004607 Human gen
C 42	62.5	3.1	156	6	CQ397944	CQ397944 Sequence
C 43	62.5	3.1	156	6	CQ404236	CQ404236 Sequence
C 44	62	3.0	185	10	MMASPART5	X07306 Mouse aspar
C 45	62	3.0	188	6	BD071586	BD071586 Secreted

ALIGNMENTS

RESULT 1
CQ110706/c
LOCUS CQ110706 180 bp DNA linear PAT 21-JAN-2004
DEFINITION Sequence 19565 from Patent WO0157272.
ACCESSION CQ110706
VERSION CQ110706.1 GI:41080089
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
TITLE Human genome-derived single exon nucleic acid probes useful for
JOURNAL analysis of gene expression in human placenta
PATENT: WO 0157272-A 19565 09-AUG-2001,
Acemica, Inc. (US)
FEATURES
LOCATION/Qualifiers
SOURCE
1..180
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC006501.5-EXPRESSED IN PLACENTA. SIGNAL =
1.9-BEST_HUMAN HIT: AL040518.1, EVALU8 1.00E-94-NT HIT:
U43139.1, EVALU8 1.10E-01-SWISSPROT HIT: Q16760, EVALU8
6.80E-02"

ORIGIN

Alignment Scores:
Pred. No.: 6.53e-19 Length: 180
Score: 300.00 Matches: 58
Percent Similarity: 98.31% Conservative: 0
Best Local Similarity: 98.31% Mismatches: 1
Query Match: 14.66% Indels: 0
Gaps: 0
DB: 6

US-10-077-111-13 (1-384) x CQ110706 (1-180)

QY 227 ArgArgThGluHhGlnLeuLysGlnPheThrGluAspTrpSerGluGluValSer 246
178 AGCGGCAACAGACATCGCTGAAGCAATTACGAAATGCTCAGAGATGCTCA 119

QY 247 ThrTrpLeuCyAlaGlnAspLeuLysAspLeuValGlyIlePheLysMetAsnAsnIle 266
118 ACATGGCTTTGTGCACAAAGATTAAAGATCTTGCTGATTTTCAAGATGAATACATT 59

QY 267 AspGlyLysGlnLeuLeuAsnLeuThrLysGluSerLeuAlaAspAspLeuLysIle 285
58 GATGAAAGAAAGACTGTTGAATCTTACAAAGAAAGTCTGCTGATGATTTGAAAAATT 2

RESULT 2
CQ149443/c
LOCUS CQ149443 180 bp DNA linear PAT 21-JAN-2004
DEFINITION Sequence 19465 from Patent WO0157276.
ACCESSION CQ149443
VERSION CQ149443.1 GI:41156793
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
TITLE Human genome-derived single exon nucleic acid probes useful for
JOURNAL analysis of gene expression in human bone marrow
Patent: WO 0157276-A 19465 09-AUG-2001;
Neomica, Inc. (US)
FEATURES
Source location/Qualifiers
1..180
/organism="Homo sapiens"
/mol_type="unasigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC006501.5-EXPRESSED IN BONE MARROW, SIGNAL
= 0.81-BEST HUMAN HIT: AL040518.1, EVALUE 1.00e-94-NT HIT:
U43139.1, EVALUE 1.10e-01-SWISSPROT HIT: Q16760, EVALUE
6.80e-02"

ORIGIN
Alignment Scores:
Pred. No.: 6.53e-19 Length: 180
Score: 300.00 Matches: 58
Percent Similarity: 98.31% Conservative: 0
Best Local Similarity: 98.31% Mismatches: 1
Query Match: 14.66% Indels: 0
Gaps: 0
DB: 6

US-10-077-111-13 (1-384) x CQ149443 (1-180)

QY 227 ArgArgThGluHhGlnLeuLysGlnPheThrGluAspTrpSerGluGluValSer 246
178 AGCGGCAACAGACATCGCTGAAGCAATTACGAAATGCTCAGAGATGCTCA 119

QY 247 ThrTrpLeuCyAlaGlnAspLeuLysAspLeuValGlyIlePheLysMetAsnAsnIle 266
118 ACATGGCTTTGTGCACAAAGATTAAAGATCTTGCTGATTTTCAAGATGAATACATT 59

QY 267 AspGlyLysGlnLeuLeuAsnLeuThrLysGluSerLeuAlaAspAspLeuLysIle 285
58 GATGAAAGAAAGACTGTTGAATCTTACAAAGAAAGTCTGCTGATGATTTGAAAAATT 2

RESULT 3

CQ232723/c
LOCUS CQ232723 180 bp DNA linear PAT 21-JAN-2004
DEFINITION Sequence 19562 from Patent WO0157273.
ACCESSION CQ232723
VERSION CQ232723.1 GI:41215941
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
TITLE HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
JOURNAL ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER<130> PB 0004 WO
3<150> US 60/180,312<151> 04 February 2000 (04.02.00)<150> US
60/207,456<151> 26 May 2000 (26.05.00)<150> US 09/632,366<151> 03
August 2000 (03.08.00)<150> GB 24263,6<151> 03 October 2000
(03.10.00)<150> US 60/236,359<151> 27 September 2000
(27.09.00)<150> US 60/234,687<151> 21 September 2000
(21.09.00)<150> US 09/608,408<151> 30 June 2000 (30.06.00)<170>
Molecular Dynamics Sequence Labeling Engine
Patent: WO 0157273-A 19562 09-AUG-2001;
Neomica, Inc. (US)
FEATURES
Source location/Qualifiers
1..180
/organism="Homo sapiens"
/mol_type="unasigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC006501.5-EXPRESSED IN ADULT LIVER, SIGNAL
= 0.72-BEST HUMAN HIT: AL040518.1, EVALUE 1.00e-94-NT HIT:
U43139.1, EVALUE 1.10e-01-SWISSPROT HIT: Q16760, EVALUE
6.80e-02"

ORIGIN
Alignment Scores:
Pred. No.: 6.53e-19 Length: 180
Score: 300.00 Matches: 58
Percent Similarity: 98.31% Conservative: 0
Best Local Similarity: 98.31% Mismatches: 1
Query Match: 14.66% Indels: 0
Gaps: 0
DB: 6

US-10-077-111-13 (1-384) x CQ232723 (1-180)

QY 227 ArgArgThGluHhGlnLeuLysGlnPheThrGluAspTrpSerGluGluValSer 246
178 AGCGGCAACAGACATCGCTGAAGCAATTACGAAATGCTCAGAGATGCTCA 119

QY 247 ThrTrpLeuCyAlaGlnAspLeuLysAspLeuValGlyIlePheLysMetAsnAsnIle 266
118 ACATGGCTTTGTGCACAAAGATTAAAGATCTTGCTGATTTTCAAGATGAATACATT 59

QY 267 AspGlyLysGlnLeuLeuAsnLeuThrLysGluSerLeuAlaAspAspLeuLysIle 285
58 GATGAAAGAAAGACTGTTGAATCTTACAAAGAAAGTCTGCTGATGATTTGAAAAATT 2

RESULT 4
CQ270754/c
LOCUS CQ270754 180 bp DNA linear PAT 23-JAN-2004
DEFINITION Sequence 19015 from Patent WO0157277.
ACCESSION CQ270754
VERSION CQ270754.1 GI:41243358
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
TITLE Human genome-derived single exon nucleic acid probes useful for
JOURNAL analysis of gene expression in human fetal liver
Patent: WO 0157277-A 19015 09-AUG-2001;
Neomica, Inc. (US)

FEATURES
Source

Location/Qualifiers
1. 180
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC006501.5-EXPRESSED IN FETAL LIVER, SIGNAL = 1.3-EST_HUMAN HIT: AL040518.1, EVALUE 1.00e-94-NT HIT: U43139.1, EVALUE 1.10e-01-SWISSPROT HIT: Q16760, EVALUE 6.80e-02"

ORIGIN

Alignment Scores:

Pred. No.: 6.53e-19 Length: 180
Score: 300.00 Matches: 58
Percent Similarity: 98.31% Conservative: 0
Best Local Similarity: 98.31% Mismatches: 1
Query Match: 14.66% Indels: 0
DB: 6 Gaps: 0

US-10-077-111-13 (1-384) x CQ270754 (1-180)

Qy 227 ArgArGrThrGluHisGlnLeuLysGlnPheThrGluAspTrpSerGluGluValSer 246

Db 178 AGGCGCAGACAACATCGCTGAAGCAATTACCGAAGATTGTCAGAGAGATGCTCA 119

Qy 247 ThrTrpLeuCysAlaGlnAspLeuLysAspLeuValGlyIlePheLysMetAsnAsnIle 266

Db 118 ACATGGCTTTGTGCACAAAGATTTTAAAGATCTTTGGTATTTTCAAGATGATAATTAATT 59

Qy 267 AspGlyLysGluLeuLeuAsnLeuThrLysGluSerLeuAlaAspAspLeuLysIle 285

Db 58 GATGAGAAAGAACTGTGAATCTTACAAAAGAAAGTGTGCTGATGATTTGAAAATT 2

RESULT 5

CQ308036/c CQ308036 180 bp DNA linear PAT 23-JAN-2004

LOCUS Sequence 19141 from Patent WO0186003.

ACCESSION CQ308036

VERSION CQ308036.1 GI:41268613

KEYWORDS

ORGANISM

SOURCE

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

Source

1. 180

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

/note="MAP TO AC006501.5-EXPRESSED IN LUNG, SIGNAL =

0.88-EST_HUMAN HIT: AL040518.1, EVALUE 1.00e-94-NT HIT: U43139.1, EVALUE 1.10e-01-SWISSPROT HIT: Q16760, EVALUE 6.80e-02"

ORIGIN

Alignment Scores:

Pred. No.: 6.53e-19 Length: 180
Score: 300.00 Matches: 58
Percent Similarity: 98.31% Conservative: 0
Best Local Similarity: 98.31% Mismatches: 1
Query Match: 14.66% Indels: 0
DB: 6 Gaps: 0

US-10-077-111-13 (1-384) x CQ308036 (1-180)

Qy 227 ArgArGrThrGluHisGlnLeuLysGlnPheThrGluAspTrpSerGluGluValSer 246

Db 178 AGGCGCAGACAACATCGCTGAAGCAATTACCGAAGATTGTCAGAGAGATGCTCA 119

Qy 247 ThrTrpLeuCysAlaGlnAspLeuLysAspLeuValGlyIlePheLysMetAsnAsnIle 266

Db 118 ACATGGCTTTGTGCACAAAGATTTTAAAGATCTTTGGTATTTTCAAGATGATAATTAATT 59

Db 178 AGGCGCAGACAACATCGCTGAAGCAATTACCGAAGATTGTCAGAGAGATGCTCA 119

Qy 247 ThrTrpLeuCysAlaGlnAspLeuLysAspLeuValGlyIlePheLysMetAsnAsnIle 266

Db 118 ACATGGCTTTGTGCACAAAGATTTTAAAGATCTTTGGTATTTTCAAGATGATAATTAATT 59

Qy 267 AspGlyLysGluLeuLeuAsnLeuThrLysGluSerLeuAlaAspAspLeuLysIle 285

Db 58 GATGAGAAAGAACTGTGAATCTTACAAAAGAAAGTGTGCTGATGATTTGAAAATT 2

RESULT 6

CQ344856/c CQ344856 180 bp DNA linear PAT 23-JAN-2004

LOCUS Sequence 18950 from Patent WO0157275.

DEFINITION CQ344856

ACCESSION CQ344856

VERSION CQ344856.1 GI:41293927

KEYWORDS

ORGANISM

SOURCE

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

Source

1. 180

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

/note="MAP TO AC006501.5-EXPRESSED IN BRAIN, SIGNAL =

0.73-EST_HUMAN HIT: AL040518.1, EVALUE 1.00e-94-NT HIT: U43139.1, EVALUE 1.10e-01-SWISSPROT HIT: Q16760, EVALUE 6.80e-02"

US-10-077-111-13 (1-384) x CQ344856 (1-180)

Qy 227 ArgArGrThrGluHisGlnLeuLysGlnPheThrGluAspTrpSerGluGluValSer 246

Db 178 AGGCGCAGACAACATCGCTGAAGCAATTACCGAAGATTGTCAGAGAGATGCTCA 119

Qy 247 ThrTrpLeuCysAlaGlnAspLeuLysAspLeuValGlyIlePheLysMetAsnAsnIle 266

Db 118 ACATGGCTTTGTGCACAAAGATTTTAAAGATCTTTGGTATTTTCAAGATGATAATTAATT 59

Qy 267 AspGlyLysGluLeuLeuAsnLeuThrLysGluSerLeuAlaAspAspLeuLysIle 285

Db 58 GATGAGAAAGAACTGTGAATCTTACAAAAGAAAGTGTGCTGATGATTTGAAAATT 2

RESULT 7

CQ058095/c CQ058095 132 bp DNA linear PAT 19-JAN-2004

LOCUS Sequence 8915 from Patent WO0157270.

DEFINITION CQ058095

ACCESSION CQ058095.1 GI:41032601

VERSION CQ058095.1

KEYWORDS

ORGANISM

SOURCE

Homo sapiens (human)

0.73-EST_HUMAN HIT: AL040518.1, EVALUE 1.00e-94-NT HIT: U43139.1, EVALUE 1.10e-01-SWISSPROT HIT: Q16760, EVALUE 6.80e-02"

ORIGIN

Alignment Scores:

Pred. No.: 6.53e-19 Length: 180
Score: 300.00 Matches: 58
Percent Similarity: 98.31% Conservative: 0
Best Local Similarity: 98.31% Mismatches: 1
Query Match: 14.66% Indels: 0
DB: 6 Gaps: 0

US-10-077-111-13 (1-384) x CQ344856 (1-180)

Qy 227 ArgArGrThrGluHisGlnLeuLysGlnPheThrGluAspTrpSerGluGluValSer 246

Db 178 AGGCGCAGACAACATCGCTGAAGCAATTACCGAAGATTGTCAGAGAGATGCTCA 119

Qy 247 ThrTrpLeuCysAlaGlnAspLeuLysAspLeuValGlyIlePheLysMetAsnAsnIle 266

Db 118 ACATGGCTTTGTGCACAAAGATTTTAAAGATCTTTGGTATTTTCAAGATGATAATTAATT 59

TITLE Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human breast and hbl 100 cells
JOURNAL Patent: WO 0157270-A 8915 09-AUG-2001;
Neomica, Inc. (US)
FEATURES location/Qualifiers
source 1..132
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC009307.1-EXPRESSED IN HBL100, SIGNAL = 0.68-EST HUMAN HIT: BE675766.1, EVALUE 2.00e-68-NT HIT: X71133.1, EVALUE 1.20e+00-SWISSPROT HIT: P54860, EVALUE 1.00e+00"

ORIGIN

Alignment Scores:

Pred. No.:	8.22e-11	Length:	132
Score:	210.00	Matches:	43
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.26%	Indels:	0
DB:	6	Gaps:	0

US-10-077-111-13 (1-384) x CQ058095 (1-132)

QY 290 LeuArgSerLyValLeuArgLyIleGluGluLeuArgThrLyValLySerLeuSer 309
| | | | |
Db 130 CTCGCTAGTAAGTCTGAGGAAATTGAAAGAGCTCAGACCAAGTTAAATCCCTTCT 71
| | | | |
QY 310 SerGlyIleProAspGluPheIleCysProIleThrArgGluLeuMetLyAspProVal 329
| | | | |
Db 70 TCAGGAATTCCTGATGAAATTATATATGCAATACCTAGAGAACTTATGAAAGATCCGGTC 11
| | | | |

QY 330 IleAlaSer 332
| | | | |
Db 10 ATCGCATCA 2

RESULT 8
CQ077416/c CQ077416 132 bp DNA linear PAT 20-JAN-2004
LOCUS Sequence 13216 from Patent WO0157278.
DEFINITION CQ077416
ACCESSION CQ077416
VERSION CQ077416.1 GI:41047285
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
AUTHORS Human genome-derived single exon nucleic acid probes useful for
TITLE analysis of gene expression in human hela cells or other human
cervical epithelial cells

JOURNAL Patent: WO 0157278-A 13216 09-AUG-2001;
Neomica, Inc. (US)
FEATURES location/Qualifiers
source 1..132
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC009307.1-EXPRESSED IN HELA, SIGNAL = 0.89-EST HUMAN HIT: BE675766.1, EVALUE 2.00e-68-NT HIT: X71133.1, EVALUE 1.20e+00-SWISSPROT HIT: P54860, EVALUE 1.00e+00"

ORIGIN

Alignment Scores:

Pred. No.:	8.22e-11	Length:	132
Score:	210.00	Matches:	43
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.26%	Indels:	0
DB:	6	Gaps:	0

US-10-077-111-13 (1-384) x CQ077416 (1-132)

QY 290 LeuArgSerLyValLeuArgLyIleGluGluLeuArgThrLyValLySerLeuSer 309
| | | | |
Db 130 CTCGCTAGTAAGTCTGAGGAAATTGAAAGAGCTCAGACCAAGTTAAATCCCTTCT 71
| | | | |
QY 310 SerGlyIleProAspGluPheIleCysProIleThrArgGluLeuMetLyAspProVal 329
| | | | |
Db 70 TCAGGAATTCCTGATGAAATTATATATGCAATACCTAGAGAACTTATGAAAGATCCGGTC 11
| | | | |

QY 330 IleAlaSer 332
| | | | |
Db 10 ATCGCATCA 2

RESULT 9
CQ108429/c CQ108429 132 bp DNA linear PAT 21-JAN-2004
LOCUS Sequence 17288 from Patent WO0157272.
DEFINITION CQ108429
ACCESSION CQ108429
VERSION CQ108429.1 GI:41077482
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
AUTHORS Human genome-derived single exon nucleic acid probes useful for
TITLE analysis of gene expression in human placenta
JOURNAL Patent: WO 0157272-A 17288 09-AUG-2001;
Neomica, Inc. (US)
FEATURES location/Qualifiers
source 1..132
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC009307.1-EXPRESSED IN PLACENTA, SIGNAL = 0.79-EST HUMAN HIT: BE675766.1, EVALUE 2.00e-68-NT HIT: X71133.1, EVALUE 1.20e+00-SWISSPROT HIT: P54860, EVALUE 1.00e+00"

ORIGIN

Alignment Scores:

Pred. No.:	8.22e-11	Length:	132
Score:	210.00	Matches:	43
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.26%	Indels:	0
DB:	6	Gaps:	0

US-10-077-111-13 (1-384) x CQ108429 (1-132)

QY 290 LeuArgSerLyValLeuArgLyIleGluGluLeuArgThrLyValLySerLeuSer 309
| | | | |
Db 130 CTCGCTAGTAAGTCTGAGGAAATTGAAAGAGCTCAGACCAAGTTAAATCCCTTCT 71
| | | | |
QY 310 SerGlyIleProAspGluPheIleCysProIleThrArgGluLeuMetLyAspProVal 329
| | | | |
Db 70 TCAGGAATTCCTGATGAAATTATATGCAATACCTAGAGAACTTATGAAAGATCCGGTC 11
| | | | |

QY 330 IleAlaSer 332
| | | | |
Db 10 ATCGCATCA 2

RESULT 10
CQ147070/c CQ147070 132 bp DNA linear PAT 21-JAN-2004
LOCUS Sequence 17092 from Patent WO0157276.
DEFINITION CQ147070
ACCESSION CQ147070
VERSION CQ147070.1 GI:41154420
KEYWORDS
SOURCE Homo sapiens (human)

DB 10 ATCGCATCA 2
|||||
RESULT 13
LOCUS CQ268431/c 132 bp DNA linear PAT 23-JAN-2004
DEFINITION Sequence 16692 from Patent WO0157277.
ACCESSION CQ268431
VERSION CQ268431.1 GI:41241035
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
TITLE Human genome-derived single exon nucleic acid probes useful for
analysis of gene expression in human fetal liver
JOURNAL Patent: WO 0157277-A 1692 09-AUG-2001;
Neomica, Inc. (US)
location/Qualifiers
1..132
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC009307.1-EXPRESSED IN FETAL LIVER, SIGNAL
= 0.78-EST HUMAN HIT: BE675766.1, EVALUE 2.00e-68-NT HIT:
X71133.1, EVALUE 1.20e+00-SWISSPROT HIT: P54860, EVALUE
1.00e+00"
ORIGIN
Alignment Scores:
Pred. No.: 8.22e-11 Length: 132
Score: 210.00 Matches: 43
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.26% Indels: 0
DB: 6 Gaps: 0
US-10-077-111-13 (1-384) x CQ268431 (1-132)
QY 290 LeuArgSerLyValLeuArgLySllEgLIuLeuArgThryVallySerLeuSer 309
DB 130 CTCGCTAGTAAAGTCTGAGGAAATTGAAGAGCTCGAGCAAGCTTAATCCCTTCT 71
QY 310 SerGlyIleProAspGluPheIleCysProIleThrArgGluLeuMerLysAspProVal 329
DB 70 TCAGGAATTCCTGATGAAATTATATGTCCAATACCTAGAACTTATGAAGAATCCGCTC 11
QY 330 IleAlaSer 332
DB 10 ATCGCATCA 2
RESULT 14
LOCUS CQ305465/c 132 bp DNA linear PAT 23-JAN-2004
DEFINITION Sequence 16570 from Patent WO0186003.
ACCESSION CQ305465
VERSION CQ305465.1 GI:41266042
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
TITLE Human genome-derived single exon nucleic acid probes useful for
analysis of gene expression in human lung
JOURNAL Patent: WO 0186003-A 16570 15-NOV-2001;
Neomica, Inc. (US)
location/Qualifiers
1..132
/organism="Homo sapiens"
/db_xref="taxon:9606"

/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC009307.1-EXPRESSED IN LUNG, SIGNAL =
0.76-EST HUMAN HIT: BE675766.1, EVALUE 2.00e-68-NT HIT:
X71133.1, EVALUE 1.20e+00-SWISSPROT HIT: P54860, EVALUE
1.00e+00"
ORIGIN
Alignment Scores:
Pred. No.: 8.22e-11 Length: 132
Score: 210.00 Matches: 43
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.26% Indels: 0
DB: 6 Gaps: 0
US-10-077-111-13 (1-384) x CQ305465 (1-132)
QY 290 LeuArgSerLyValLeuArgLySllEgLIuLeuArgThryVallySerLeuSer 309
DB 130 CTCGCTAGTAAAGTCTGAGGAAATTGAAGAGCTCGAGCAAGCTTAATCCCTTCT 71
QY 310 SerGlyIleProAspGluPheIleCysProIleThrArgGluLeuMerLysAspProVal 329
DB 70 TCAGGAATTCCTGATGAAATTATATGTCCAATACCTAGAACTTATGAAGAATCCGCTC 11
QY 330 IleAlaSer 332
DB 10 ATCGCATCA 2
RESULT 15
LOCUS AX905781 191 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 21644 from Patent EP1033401.
ACCESSION AX905781
VERSION AX905781.1 GI:40060762
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Dumais, Milne, Edwards, J.B., Duclet, A. and Giordano, J.Y.
TITLE Expressed sequence tags and encoded human proteins
JOURNAL Patent: EP 1033401-A 21644 06-SEP-2000;
Genetec (FR)
location/Qualifiers
1..191
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 42.4 Length: 191
Score: 85.00 Matches: 18
Percent Similarity: 50.00% Conservative: 7
Best Local Similarity: 36.00% Mismatches: 25
Query Match: 4.15% Indels: 0
DB: 6 Gaps: 0
US-10-077-111-13 (1-384) x AX905781 (1-191)
QY 100 ValArgValCysGlnInheSerProAspSerThrCysLeuAlaSerGlyAlaAlaAspGly 119
DB 42 GTTACTGCTGCTCCATCTCCAGAGCTGACAGCACTCTGCTGACCTGAGAGAG 101
QY 120 ThrValIleuThrPheAlaGlnSerTyLyLeuTyArgCysGlySerVallyAsp 139
DB 102 TCGGCTCTTCTATGAGCACTAGAGCTCTACAGTMAATTCCGAGACTAGAGGCCATCA 161
QY 140 GlySerLeuAlaIaCysAlaPheSerPro 149
|||||

Db 162 AGCAGTGTGTCCTTGACTTCCCCC 191

Search completed: February 5, 2005, 16:54:50
Job time : 5696 secs

Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: February 5, 2005, 12:08:02 ; Search time 698 Seconds
(without alignments)
3256.706 Million cell updates/sec

Title: US-10-077-111-13
Perfect score: 2047
Sequence: 1 MVKLIHTLADHGDDVNCACF.....LTPNRLKMAINRWLETHQK 384

Scoring table: BIOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Delop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 segs, 2959870667 residues
Total number of hits satisfying chosen parameters: 4887668

Minimum DB seg length: 0
Maximum DB seg length: 200

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+.p2n.model -DEV=x1h
-Q=/cgr2_1/USPTO.spool/US10077111/runat_04022005_121140_7174/app_query.fasta_1.583
-DB=N_Geneseq_16Dec04 -OPMT=fastcap -SUFFIX=p2nslm200.rng -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bases -START=1 -END=-1 -MATRIX=blcosum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=200 -USER=US10077111_@CGN_1_1_644_@runat_04022005_121140_7174 -NCPV=6
-ICPV=3 -NO_MMAPP -LARGEOQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONCLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	300	14.7	180	4	ABA70710 Human foe
C 2	300	14.7	180	4	AA150879 Probe #19
C 3	300	14.7	180	4	AA444908 Human bra
C 4	300	14.7	180	4	AAK18959 Human bra
C 5	300	14.7	180	4	ABA44572 Human liv

C 6	300	14.7	180	6	ABS19150 Human gen
C 7	210	10.3	132	4	AA123283 Probe #13
C 8	210	10.3	132	4	ABA68387 Human foe
C 9	210	10.3	132	4	AA148602 Probe #17
C 10	210	10.3	132	4	ABA50437 Human bre
C 11	210	10.3	132	4	AAK42535 Human bon
C 12	210	10.3	132	4	ABS42143 Human liv
C 13	210	10.3	132	5	AA108924 Probe #89
C 14	210	10.3	132	6	ABS16579 Human gen
C 15	150	7.3	92	4	AAH36289 Human col
C 16	93.5	4.6	174	12	ACH93376 Human gen
C 17	87	4.3	150	12	ACH85482 Human gen
C 18	85	4.2	191	3	AA17569 Human sec
C 19	82	4.0	173	4	ABA74172 Human foe
C 20	82	4.0	173	4	AA154624 Probe #23
C 21	82	4.0	173	4	AAK48795 Human bon
C 22	82	4.0	173	4	AAK22627 Human bra
C 23	82	4.0	173	4	ABS48468 Human liv
C 24	82	4.0	173	6	ABS22500 Human gen
C 25	79.5	3.9	169	10	ACA55819 C. elegans
C 26	79.5	3.9	159	12	AD155615 Human pol
C 27	78.5	3.8	155	12	ACH82916 Human gen
C 28	73	3.6	190	2	AAV11599 Homo sapi
C 29	73	3.6	190	5	AA198377 Human CDN
C 30	72.5	3.5	197	6	ABV89016 Human col
C 31	72.5	3.5	197	6	ABV66894 Human col
C 32	67.5	3.5	41	4	AA164919 Beta-cran
C 33	67.5	3.3	131	12	ACH92306 Human gen
C 34	66	3.2	167	10	ABX82592 Corn ear-
C 35	65.5	3.2	154	6	ABV96402 Human pan
C 36	65	3.2	131	4	AA126039 Probe #15
C 37	65	3.2	131	4	ABA73203 Human foe
C 38	65	3.2	131	4	AA153631 Probe #22
C 39	65	3.2	131	4	ABA48636 Probe #17
C 40	65	3.2	131	4	AAK47803 Human bon
C 41	65	3.2	131	4	AAK21638 Human bra
C 42	65	3.2	131	4	ABS47528 Human liv
C 43	65	3.2	111	6	ABS21810 Human gen
C 44	64	3.1	118	12	ACH83189 Human gen
C 45	64	3.1	142	12	ACH83110 Human gen

ALIGNMENTS

RESULT 1
ID ABA70710 standard; DNA; 180 BP.
XX
AC ABA70710;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human foetal liver single exon nucleic acid probe #19015.
XX
KM Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000669.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLB-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2001-483447/52.
XX
DR
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human fetal liver.
XX
PS Claim 4; SEQ ID NO 19015; 6339p + Sequence Listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human foetal liver. The
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human fetal liver. The
CC present sequence is a single exon nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at fcp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 180 BP; 49 A; 42 C; 25 G; 64 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 1,15e-20 Length: 180
Score: 300.00 Matches: 58
Percent Similarity: 98.31% Conservative: 0
Best Local Similarity: 98.31% Mismatches: 1
Query Match: 14.66% Indels: 0
DB: Gaps: 0
US-10-077-111-13 (1-384) x ABA70710 (1-180)
QY 227 ArgArgThrGluHisGlnLeuLysGlnPheThrGluAspTrpSerGluGluValSer 246
Db 178 AGCGCACAGAACATCAGTGAAGCAATTACCGAATTGTCAGAGAGATGCTCA 119
QY 247 ThTTpLeuCYeAlaGlnAspLeuLysAspLeuValGlyIlePheLysMetAsnAsnIle 266
Db 118 ACATGGCTTTGTGCACAAAGATTAAAGATCTTGTGGTATTTTCAAGATGAATACATT 59
QY 267 AspGlyLYeGlnLeuLeuAsnLeuThrLYeGlnSerLeuAlaAspAspLeuLysIle 285
Db 58 GATGAAAAGAACTGTTGAATCTTACAAAGAAAGTGGCTGATGATTTGAAATTT 2
RESULT 2
ID AAI50879/c
XX AAI50879 standard; DNA; 180 BP.
XX
AC AAI50879;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #19565 used to measure gene expression in human placenta sample.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN W0200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000663.
XX
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX

XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2001-48897/53.
XX
DR
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX
PS Claim 25; SEQ ID NO 19565; 6549p; English.
XX
CC The present invention relates to single exon nucleic acid probes (SNP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders
XX
SQ Sequence 180 BP; 49 A; 42 C; 25 G; 64 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 1,15e-20 Length: 180
Score: 300.00 Matches: 58
Percent Similarity: 98.31% Conservative: 0
Best Local Similarity: 98.31% Mismatches: 1
Query Match: 14.66% Indels: 0
DB: Gaps: 0
US-10-077-111-13 (1-384) x AAI50879 (1-180)
QY 227 ArgArgThrGluHisGlnLeuLysGlnPheThrGluAspTrpSerGluGluValSer 246
Db 178 AGCGCACAGAACATCAGTGAAGCAATTACCGAATTGTCAGAGAGATGCTCA 119
QY 247 ThTTpLeuCYeAlaGlnAspLeuLysAspLeuValGlyIlePheLysMetAsnAsnIle 266
Db 118 ACATGGCTTTGTGCACAAAGATTAAAGATCTTGTGGTATTTTCAAGATGAATACATT 59
QY 267 AspGlyLYeGlnLeuLeuAsnLeuThrLYeGlnSerLeuAlaAspAspLeuLysIle 285
Db 58 GATGAAAAGAACTGTTGAATCTTACAAAGAAAGTGGCTGATGATTTGAAATTT 2
RESULT 3
ID AAK44908/c
XX AAK44908 standard; DNA; 180 BP.
XX
AC AAK44908;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 19465.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
XX microarray; cancer; leukemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.
XX
PN W0200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000668.
XX
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX

DR WPI; 2001-488900/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human bone marrow.
XX
PS Example 4; SEQ ID NO 19465; 658bp + Sequence listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukemia and myeloma. The present sequence is one of
CC the probes of the invention
XX
SQ Sequence 180 BP; 49 A; 42 C; 25 G; 64 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 1,15e-20 Length: 180
Score: 300.00 Matches: 58
Percent Similarity: 98.31% Conservative: 0
Best Local Similarity: 98.31% Mismatches: 1
Query Match: 14.66% Indels: 0
DB: 4 Gaps: 0
US-10-077-111-13 (1-384) x AAK44908 (1-180)
QY 227 ArgArGThrGluHisGlnLeuLysGlnPheThrGluAspTrpSerGluIuValSer 246
DB 178 AGGGGCACAGAACATCAGCTGAGCAATTACCGAAGATTGGTCAGAGAGATGCTCA 119
QY 247 ThrTrpLeuCySAIaGlnAspLeuLysAspLeuValGlyIlePheLysMetAsnAsnIle 266
DB 118 ACATGGCTTTGTGCACAAAGATTAAAGATCTTGCTATTTCACAGATGAATTAACATT 59
QY 267 AspGlyLysGlnLeuLeuAsnLeuThrLysGlnSerLeuAlaAspAspLeuLysIle 285
DB 58 GATGGAAGAAGAACTGTTGAATCTTACAAAGAAAGTCTGCTGATGATTGAAATTT 2
RESULT 4
AAK18959/c
ID AAK18959 standard; DNA; 180 BP.
XX
XX AAK18959;
XX AC
XX DT 05-NOV-2001 (first entry)
XX DB Human brain expressed single exon probe SEQ ID NO: 18950.
XX DE Human brain expressed single exon probe SEQ ID NO: 18950.
XX KW Human; brain expressed exon; gene expression analysis; probe; microarray;
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
KW ss.
XX
XX Homo sapiens.
XX OS
XX PN WO200157275-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000667.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX PA
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483446/52.
DR

XX
XX Single exon nucleic acid probes for analyzing gene expression in human
PT brain.
XX
XX
PS Example 4; SEQ ID NO 18950; 650bp + Sequence listing; English.
XX
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention
XX
SQ Sequence 180 BP; 49 A; 42 C; 25 G; 64 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 1,15e-20 Length: 180
Score: 300.00 Matches: 58
Percent Similarity: 98.31% Conservative: 0
Best Local Similarity: 98.31% Mismatches: 1
Query Match: 14.66% Indels: 0
DB: 4 Gaps: 0
US-10-077-111-13 (1-384) x AAK18959 (1-180)
QY 227 ArgArGThrGluHisGlnLeuLysGlnPheThrGluAspTrpSerGluIuValSer 246
DB 178 AGGGGCACAGAACATCAGCTGAGCAATTACCGAAGATTGGTCAGAGAGATGCTCA 119
QY 247 ThrTrpLeuCySAIaGlnAspLeuLysAspLeuValGlyIlePheLysMetAsnAsnIle 266
DB 118 ACATGGCTTTGTGCACAAAGATTAAAGATCTTGCTATTTCACAGATGAATTAACATT 59
QY 267 AspGlyLysGlnLeuLeuAsnLeuThrLysGlnSerLeuAlaAspAspLeuLysIle 285
DB 58 GATGGAAGAAGAACTGTTGAATCTTACAAAGAAAGTCTGCTGATGATTGAAATTT 2
RESULT 5
ABS44572/c
ID ABS44572 standard; DNA; 180 BP.
XX
XX ABS44572;
XX AC
XX DT 25-FEB-2003 (first entry)
XX DB Human liver single exon probe, SEQ ID NO 19562.
XX DE Human liver single exon probe, SEQ ID NO 19562.
XX KW Human; single exon nucleic acid probe; liver; cirrhosis;
KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW coronary heart disease; ss.
XX
XX Homo sapiens.
XX OS
XX PN WO200157273-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000664.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX PA
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488898/53.
DR

XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human adult liver.
XX
XX Claim 4; SEQ ID NO 19562; 658bp; English.
XX
CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridizes at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (I) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABS25011-ABS51005 represent human
CC liver single exon nucleic acid probes of the invention. Note: The
CC sequence information for this patent does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 180 BP; 49 A; 42 C; 25 G; 64 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 1.15e-20 Length: 180
Score: 300.00 Matches: 58
Percent Similarity: 98.31% Conservative: 0
Best Local Similarity: 98.31% Mismatches: 1
Query Match: 14.66% Indels: 0
DB: Gaps: 0
US-10-077-111-13 (1-384) x ABS44572 (1-180)
QY ArgArGThGluHISGlnLeuLYSGlnPheThGluAspTrpSerGluGluValSer 246
DB 178 AGCGGCACAGAACATCAGCTGAAACCAATTACCAAGATTGGTCAGAGGAGATGTCTCA 119
QY 247 ThTTrpLeuCYsaIaGlnAspLeuLYsAspLeuValGlyIlePheLYsMeAsnAsnIle 266
DB 118 ACATGGCTTGTGCACAGATTAAAGATCTGTGATTTTCAAGATGAAATTAACATT 59
QY 267 AAGGlyLYsGlnLeuLeuAsnLeuThLYsGlnSerLeuAlaAspAspLeuLYsIle 205
DB 58 GATGGAAAGAACGTTGATCTTAACAAAGAAAGTCTGCTGATGATTTGAAAAATT 2
RESULT 6
ABS19150/C
ID ABS19150 standard; DNA; 180 BP.
XX
AC ABS19150;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human genome-derived single exon probe ORF from lung SEQ ID No 19141.
XX
XX Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
XX chronic obstructive pulmonary disease; interstitial lung disease;
XX familial idiopathic pulmonary fibrosis; neurofibromatosis;
XX tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
XX Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
XX pulmonary histiocytosis; lymphangioleiomyomatosis; Kartagener syndrome;
XX pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
XX primary ciliary dyskinesia; pulmonary hypertension;
XX hyaline membrane disease; open reading frame; ORF.
XX
OS Homo sapiens.
XX
PN WO200186003-A2.
XX
PD 15-NOV-2001.
XX
PF 30-JAN-2001; 2001WO-US000665.
XX

PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PA Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2002-114183/15.
XX
XX Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples.
XX
XX Claim 4; SEQ ID NO 19141; 634bp; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of probes
CC; the novel set of probes which hybridise at high stringency to a nucleic
CC acid expressed in the human lung; measuring gene expression in a sample
CC derived from human lung, comprising (a) contacting the array with a
CC collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of the
CC array; identifying exons in a eukaryotic genome, comprising (a)
CC algorithmically predicting at least one exon from genomic sequences of
CC the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray, assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene expression
CC analysis, and for identifying exons in a gene, particularly using human
CC lung derived mRNA and for the study of lung diseases such as asthma, lung
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
CC Kartagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a single exon probe open reading frame of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 180 BP; 49 A; 42 C; 25 G; 64 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 1.15e-20 Length: 180
Score: 300.00 Matches: 58
Percent Similarity: 98.31% Conservative: 0
Best Local Similarity: 98.31% Mismatches: 1
Query Match: 14.66% Indels: 0
DB: Gaps: 0
US-10-077-111-13 (1-384) x ABS19150 (1-180)
QY 227 ArgArGThGluHISGlnLeuLYSGlnPheThGluAspTrpSerGluGluValSer 246
DB 178 AGCGGCACAGAACATCAGCTGAAACCAATTACCAAGATTGGTCAGAGGAGATGTCTCA 119

Qy	247	ThTTPLeucYValIAGTlAspLeuYAspLeuValGlyIlePelysMeTAspAsnIle	266
Db	118	ACATGGCTTTGTTCACAAAGATTAAAGATCTTGTTGGTATTTCACAGATGAATTAACATT	59
Qy	267	AspGlyLysGluLeuLeuAnbLeuThryLysGluSerLeuAlaAspAspLeuLysIle	285
Db	58	GATGGAAAAAGAACCTGTTGAATCTTACAAAAGAAAGTCTGGCTGATGATTGAAATTT	2
RESULT 7			
ID	AA123283/c	AA123283 standard; DNA; 132 BP.	
XX	AA123283;		
XX	AA123283;		
DT	12-OCT-2001	(first entry)	
XX			
DE	Probe #13216 for gene expression analysis in human cervical cell sample.		
XX			
KW	Probe; human; microarray; gene expression; cervical epithelial cell;		
KW	cervical cancer; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	MO200157278-A2.		
XX			
PD	09-AUG-2001.		
XX			
PF	30-JAN-2001; 2001WO-US000670.		
XX			
PR	04-FEB-2000; 2000US-0180312P.		
PR	26-MAY-2000; 2000US-0207456P.		
PR	30-JUN-2000; 2000US-00608408.		
PR	03-AUG-2000; 2000US-00632366.		
PR	21-SEP-2000; 2000US-0234687P.		
PR	27-SEP-2000; 2000US-0236359P.		
PR	04-OCT-2000; 2000GB-00024263.		
XX			
PA	(MOLE-) MOLECULAR DYNAMICS INC.		
XX			
PI	Penn SG, Hanzel DK, Chen W, Rank DR;		
XX			
DR	WPI; 2001-488901/53.		
XX			
PT	Human genome-derived single exon nucleic acid probes useful for analyzing		
PT	gene expression in human cervical epithelial cells.		
XX			
PS	Claim 25; SEQ ID NO 13216; 487bp; English.		
XX			
CC	The present invention relates to human single exon nucleic acid probes		
CC	(SENPs). The present sequence is one such probe. The SENPs are derived		
CC	from human HeLa cells. The SENPs can be used to produce a single exon		
CC	microarray, which can be used for measuring human gene expression in a		
CC	sample derived from human cervical epithelial cells. By measuring gene		
CC	expression, the probes are therefore useful in grading and/or staging of		
CC	diseases of the cervix, notably cervical cancer. Note: The sequence data		
CC	for this patent did not form part of the printed specification, but was		
CC	obtained in electronic format directly from WIPO at		
XX	ftp.wipo.int/pub/published_pct_sequences		
XX			
SQ	Sequence 132 BP; 35 A; 29 C; 24 G; 44 T; 0 U; 0 Other;		
Alignment Scores:			
Pred. No.:	5,44e-12	Length:	132
Score:	210.00	Matches:	133
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.26%	Indels:	0
DB:	4	Gaps:	0
Qy	US-10-077-111-13 (1-384) x AA123283 (1-132)		
Qy	290 LeuArGSerLySValLeuArGlyLleGluGluLeuArGThrLySValLysSerLeuSer		309

```

Db      130 CTCGGTAGTAAAGTCTGAGGAAATTGAAAGCTCAGACCAAGTTAAATCCCTTCT 71
Qy      310 SerGlyIleProAspGluPheIleCysbRoiIleThirArgIleuMetIysAspProVal 329
Db      70 TCAGGAATTCCTGATGATTAATTATGATCCAACTAGAGAACTTATGAAAGATCCGGTC 11
Qy      330 IleAlaSer 332
Db      10 ATGCATCA 2

RESULT 8
ABA68387/c
ID      ABA68387 standard; DNA; 132 BP.
XX
XX      ABA68387;
AC
XX      01-FEB-2002 (first entry)
DT
XX      Human foetal liver single exon nucleic acid probe #16692.
DE
XX      Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
KW
XX      Homo sapiens.
OS
XX      WO200157277-A2.
XX      09-AUG-2001.
XX      30-JAN-2001; 2001WO-US000669.
PF
XX      04-FEB-2000; 2000US-0180312P.
PR      26-MAY-2000; 2000US-0207456P.
PR      30-JUN-2000; 2000US-00608408.
PR      03-AUG-2000; 2000US-00632366.
PR      21-SEP-2000; 2000US-0234687P.
PR      27-SEP-2000; 2000US-0236359P.
PR      04-OCT-2000; 2000GB-00024263.
XX
XX      (MOLE-) MOLECULAR DYNAMICS INC.
XX      Penn SG, Hanzel DK, Chen W, Rank DR;
XX      WPI; 2001-483447/52.
DR
XX      Human genome-derived single exon nucleic acid probes useful for analyzing
PT      gene expression in human fetal liver.
XX
XX      Claim 4; SEQ ID NO 16692; 639pp + Sequence Listing; English.
XX
XX      The invention relates to a single exon nucleic acid probe for measuring
CC      human gene expression in a sample derived from human foetal liver. The
CC      single exon nucleic acid probes may be used for predicting, measuring and/or
CC      displaying gene expression in samples derived from human fetal liver. The
CC      present sequence is a single exon nucleic acid probe of the invention.
CC      Note: The sequence data for this patent did not form part of the printed
CC      specification, but was obtained in electronic format directly from WIPO
CC      at ftp.wipo.int/pub/published_pct_sequences
XX
XX      Sequence 132 BP; 35 A; 29 C; 24 G; 44 T; 0 U; 0 Other;
SO
Alignment Scores:
Pred. No.:      5.44e-12      Length:      132
Score:          210.00      Matches:      43
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      10.26%      Indels:      0
DB:              4      Gaps:      0
US-10-077-111-13 (1-384) x ABA68387 (1-132)
Qy      290 LeuArgSerIlysValIeuArgIylIleGluIleuArgThirIysValIysSerIleuSer 309

```

```

Db      130 CTCGGTAGTAAAGTCTGAGGAAATTGAAAGCTCAGACCAAGTTAAATCCCTTCT 71
Qy      310 SerGlyIleProAspGluPheIleCysbRoiIleThirArgIleuMetIysAspProVal 329
Db      70 TCAGGAATTCCTGATGATTAATTATGATCCAACTAGAGAACTTATGAAAGATCCGGTC 11
Qy      330 IleAlaSer 332
Db      10 ATGCATCA 2

RESULT 8
ABA68387/c
ID      ABA68387 standard; DNA; 132 BP.
XX
XX      ABA68387;
AC
XX      01-FEB-2002 (first entry)
DT
XX      Human foetal liver single exon nucleic acid probe #16692.
DE
XX      Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
KW
XX      Homo sapiens.
OS
XX      WO200157277-A2.
XX      09-AUG-2001.
XX      30-JAN-2001; 2001WO-US000669.
PF
XX      04-FEB-2000; 2000US-0180312P.
PR      26-MAY-2000; 2000US-0207456P.
PR      30-JUN-2000; 2000US-00608408.
PR      03-AUG-2000; 2000US-00632366.
PR      21-SEP-2000; 2000US-0234687P.
PR      27-SEP-2000; 2000US-0236359P.
PR      04-OCT-2000; 2000GB-00024263.
XX
XX      (MOLE-) MOLECULAR DYNAMICS INC.
XX      Penn SG, Hanzel DK, Chen W, Rank DR;
XX      WPI; 2001-483447/52.
DR
XX      Human genome-derived single exon nucleic acid probes useful for analyzing
PT      gene expression in human fetal liver.
XX
XX      Claim 4; SEQ ID NO 16692; 639pp + Sequence Listing; English.
XX
XX      The invention relates to a single exon nucleic acid probe for measuring
CC      human gene expression in a sample derived from human foetal liver. The
CC      single exon nucleic acid probes may be used for predicting, measuring and/or
CC      displaying gene expression in samples derived from human fetal liver. The
CC      present sequence is a single exon nucleic acid probe of the invention.
CC      Note: The sequence data for this patent did not form part of the printed
CC      specification, but was obtained in electronic format directly from WIPO
CC      at ftp.wipo.int/pub/published_pct_sequences
XX
XX      Sequence 132 BP; 35 A; 29 C; 24 G; 44 T; 0 U; 0 Other;
SO
Alignment Scores:
Pred. No.:      5.44e-12      Length:      132
Score:          210.00      Matches:      43
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      10.26%      Indels:      0
DB:              4      Gaps:      0
US-10-077-111-13 (1-384) x ABA68387 (1-132)
Qy      290 LeuArgSerIlysValIeuArgIylIleGluIleuArgThirIysValIysSerIleuSer 309

```

Db 130 CTGCGTAGTAAGTCTGAGGAAATTGAAAGACTCAGACCAAGTTAATCCCTTCT 71
QY 310 SerGly11eProaspGluPheileCySProileThrArgGluLeuMetLysAspProval 329
Db 70 TCAGGAATTCCTGATGATTATATGTCCAATTAAGTACTTAAGAAATCCGGTC 11
QY 330 IleAlaser 332
Db 10 ATCGCATCA 2
RESULT 9
AA148602/C
ID AA148602 standard; DNA; 132 BP.
AC AA148602;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #17288 used to measure gene expression in human placenta sample.
XX
XX Probe; microarray; human; placenta; antenatal diagnosis;
KM genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000663.
XX
PR 04-FEB-2000; 2000US-0180312P.
XX
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488937/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human placenta.
XX
PS Claim 25; SEQ ID NO 17288; 654bp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders
XX
SQ Sequence 132 BP; 35 A; 29 C; 24 G; 44 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 5.44e-12 Length: 132
Score: 210.00 Matches: 43
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.26% Indels: 0
DB: 4 Gaps: 0
US-10-077-111-13 (1-384) x AA148602 (1-132)
QY 290 LeuArgSerIleValIleuArgLysIleGluGluLeuArgThrLysValLysSerLeuSer 309
Db 130 CTCGCTAGTAAGTCTGAGGAAATTGAAAGACTCAGACCAAGTTAATCCCTTCT 71
QY 310 SerGly11eProaspGluPheileCySProileThrArgGluLeuMetLysAspProval 329

Db 70 TCAGGAATTCCTGATGATTATATGTCCAATTAAGTACTTAAGAAATCCGGTC 11
QY 330 IleAlaser 332
Db 10 ATCGCATCA 2
RESULT 10
ABA50437/C
ID ABA50437 standard; DNA; 132 BP.
XX
AC ABA50437;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human breast cell single exon nucleic acid probe #9132.
XX
XX Human; microarray; single exon probe; gene expression; breast; disease;
KM cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157271-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000662.
XX
PR 04-FEB-2000; 2000US-0180312P.
XX
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-496933/54.
XX
PT New spatially-addressable set of single exon nucleic acid probes, useful
XX for measuring gene expression in sample derived from human breast,
XX comprises number of single exon nucleic acid probes.
XX
PS Claim 4; SEQ ID NO 9132; 327bp + Sequence Listing; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting the
CC probes with a collection of detectably labelled nucleic acids derived
CC from mRNA of human breast, and then measuring the label bound to each
CC probe of the microarray. The probes are useful for verifying the
CC expression of regions of genomic DNA predicted to encode proteins. They
CC are useful for gene discovery, and for determining predisposition and/or
CC assessing the toxicity of chemical agents on cells. The microarray of
CC this invention presents a far greater diversity of probes for measuring
CC gene expression, with far less bias than expressed sequence tag
CC microarrays. The method is suitable for rapid production of functional
CC information from genomic sequence. The present sequence is a single exon
CC nucleic acid probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 132 BP; 35 A; 29 C; 24 G; 44 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 5.44e-12 Length: 132
Score: 210.00 Matches: 43
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.26% Indels: 0
DB: 4 Gaps: 0

US-10-077-111-13 (1-384) x ABA50437 (1-132)

Qy 290 LeuArgSerIyValLeuArgIyValIleGluGluLeuArgThrIyValIySerLeuSer 309
Db 130 CTGGGTAGTAAAGCTGAGGAAATTGAAGACTCAGGACCAAGTTAAATCCCTTCT 71

Qy 310 SerGlyIleProAspGluPheIleCysProIleThrArgGluLeuMetIyAspProVal 329
Db 70 TCAGGAATTCCTGATGAAATTATATGTCCAATACTAGAGAACTTATGAAAGATCCGCTC 11

Qy 330 IleAlaSer 332
Db 10 ATCGCATCA 2

RESULT 11
AAK42535/c
ID AAK42535 standard; DNA; 132 BP.

AC AAK42535;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 17092.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000668.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human bone marrow.
XX
PS Example 4; SEQ ID NO 17092; 658bp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukemia and myeloma. The present sequence is one of
CC the probes of the invention
XX
SQ Sequence 132 BP; 35 A; 29 C; 24 G; 44 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5.44e-12 Length: 132
Score: 210.00 Matches: 43
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.26% Indels: 0

DB: 4 Gaps: 0
US-10-077-111-13 (1-384) x AAK42535 (1-132)

Qy 290 LeuArgSerIyValLeuArgIyValIleGluGluLeuArgThrIyValIySerLeuSer 309
Db 130 CTGGGTAGTAAAGCTGAGGAAATTGAAGACTCAGGACCAAGTTAAATCCCTTCT 71

Qy 310 SerGlyIleProAspGluPheIleCysProIleThrArgGluLeuMetIyAspProVal 329
Db 70 TCAGGAATTCCTGATGAAATTATATGTCCAATACTAGAGAACTTATGAAAGATCCGCTC 11

Qy 330 IleAlaSer 332
Db 10 ATCGCATCA 2

RESULT 12
ABS42143/c
ID ABS42143 standard; DNA; 132 BP.

AC ABS42143;
XX
DT 25-FEB-2003 (first entry)
XX
DE Human liver single exon probe, SEQ ID No 17133.
XX
KW Human; single exon nucleic acid probe; liver; cirrhosis;
KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW coronary heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157273-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000664.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-48898/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human adult liver.
XX
PS Claim 4; SEQ ID NO 17133; 658bp; English.
XX
CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (I) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABS25011-ABS51005 represent human
CC liver single exon nucleic acid probes of the invention. Note: The
CC sequence information for this patent does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 132 BP; 35 A; 29 C; 24 G; 44 T; 0 U; 0 Other;

XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of probes
CC; the novel set of probes which hybridise at high stringency to a nucleic
CC acid expressed in the human lung; measuring gene expression in a sample
CC derived from human lung, comprising (a) contacting the array with a
CC collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of the
CC array; identifying exons in a eukaryotic genome, comprising (a)
CC algorithmically predicting at least one exon from genomic sequences of
CC the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene expression
CC analysis, and for identifying exons in a gene, particularly using human
CC lung derived mRNA and for the study of lung diseases such as asthma, lung
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC Pudlak syndrome, sarcoidosis, pulmonary haemangioendothelioma, pulmonary
CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
CC Karsagen syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a single exon probe open reading frame of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 132 BP; 35 A; 29 C; 24 G; 44 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	5.44e-12	Length:	132
Score:	210.00	Matches:	43
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.26%	Indels:	0
DB:	6	Gaps:	0

US-10-077-111-13 (1-384) x ABS16579 (1-132)

QY 290 LeuArgSerLyValLeuArgLyIleGlulLeuArgThrLyValLySerLeuSer 309

Db 130 CTGGGTGTAAGGCTGAGAAATTCAGAGCTCAAGCAAGTTAAATCCCTTTCT 71

QY 310 SerLyIleProAaPgluPhelIleCySProIleThrArgIleuMetLyAspProVal 329

Db 70 TCAGGAATTCCTGATGAAATTAATGTCACATACTAGAGAACTTAAGAAATCCGCTC 11

QY 330 IleAlaSer 332

Db 10 ATCGCATCA 2

RESULT 15
ID AAH36289 standard; cDNA; 92 BP.
XX AAH36289;
XX 03-SEP-2001 (first entry)
DT Human colon cancer antigen encoding cDNA SEQ ID NO:3371.
DE

XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma; ss.
XX Homo sapiens.
OS
PN W0200122920-A2.
PM
XX
PD 05-Apr-2001.
XX
PF 28-SEP-2000; 2000WO-US026524.
XX
PR 29-SEP-1999; 99US-0157137P.
PR 03-NOV-1999; 99US-0163280P.
XX
PA (HMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Barash SC, Birse CE, Rosen CA;
XX
DR WPI: 2001-235357/24.
DR P-PSDB; AAG76884.
XX
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
XX useful for preventing, diagnosing and/or treating colorectal cancers.
PS Claim 1; Page 5177-5178; 9803jp; English.

XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where the
CC proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene therapy
CC and vaccine production. N and P may be used in the prevention, diagnosis
CC and treatment of diseases associated with inappropriate P expression. For
CC example, N and P may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of P by expressing inactive proteins or to
CC supplement the patient's own production of P. Additionally, N may be used
CC to produce the colon cancer-associated Ps, by inserting the nucleic acids
CC into a host cell and culturing the cell to express the proteins. N and P
CC can be used in the prevention, diagnosis and treatment of colorectal
CC carcinomas and cancers. AAH37196 to AAH37204 and AAB7789 represent
CC sequences used in the exemplification of the present invention. N.B.
CC Pages 666 to 682 and page 7053 of the sequence listing were missing at
CC time of publication, meaning no sequences are present for SEQ ID NO:1027
CC to 1052, 7921 and 7922

XX SQ Sequence 92 BP; 27 A; 9 C; 23 G; 28 T; 0 U; 5 Other;

Alignment Scores:

Pred. No.:	2.66e-06	Length:	92
Score:	150.00	Matches:	26
Percent Similarity:	92.86%	Conservative:	0
Best Local Similarity:	92.86%	Mismatches:	2
Query Match:	7.33%	Indels:	0
DB:	4	Gaps:	0

US-10-077-111-13 (1-384) x AAH36289 (1-92)

QY 161 GlYAspLeuThrValITPAspAspLyMetArgCySLeuHisSerGluValAlaHisAsp 180

Db 1 GGTGATTTAAACAGTGTGGATGATTAATAAGTGTGTGATGNGAAGCAACATGAT 60

QY 181 LeuGlyIleThrCySAspPhe 188

Db 61 CTTGGAATTAACCTGCTGNGATTTT 84

Search completed: February 5, 2005, 15:19:50
Job time : 701 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 5, 2005, 14:52:43 ; Search time 4204 Seconds
(without alignments)
3476.850 Million cell updates/sec

Title: US-10-077-111-13

Perfect score: 2047

Sequence: 1 MVXLIHTLADHDGDDVNCAP.....LTPNRLTKAKAINMLETHOK 384

Scoring table:

BLISSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 3966416

Minimum DB seq length: 0
Maximum DB seq length: 200

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10077111/runat_04022005_121141_7194/app_query.fasta.1.583
-DB=EST -OPMT=fastcap -SUFFIX=pnszlm200.rst -MINMATCH=0.1 -LOOPT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blissum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200
-USRR=US10077111 @CGN_1_1_4352 @runat_04022005_121141_7194 -NCPU=6 -ICPU=3
-NO MMAP -LARGEROVER -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_hlc: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_g981: *
9: gb_g982: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	345	16.9	199	2	BE669639 7e16e07.x
C 2	342	16.7	199	2	BF110272 7n51907.x
C 3	290	14.2	160	1	AA501581 ne97q11.s
C 4	289	14.1	199	1	A1415375 A1415375
C 5	209.5	10.2	190	2	BF653223 mc35f10.x
C 6	174	8.5	138	6	CA587431 LB812p34
C 7	151	7.4	191	2	BB323974 BB323974
C 8	146.5	7.2	144	1	A1622018 A1622018
C 9	138	6.7	179	9	CC724603 CC724603

C 10	136	6.6	82	1	AA912881
C 11	129	6.3	184	9	AG211342
C 12	122	6.0	175	2	BE240075
C 13	117	5.7	145	9	CG460443
C 14	113	5.5	189	2	AM032239
C 15	111	5.4	177	2	BB186455
C 16	108	5.3	131	9	CG460401
C 17	103.5	5.1	195	1	AI202087
C 18	99	4.8	153	5	BO760356
C 19	95.5	4.7	102	7	CR455364
C 20	88	4.3	182	7	CO261506
C 21	87.5	4.3	174	7	CNS79930
C 22	86.5	4.2	181	2	BE461636
C 23	86	4.2	132	1	AF071906
C 24	85	4.2	172	1	AF071907
C 25	85	4.2	139	7	CK617095
C 26	84.5	4.1	199	5	BW563342
C 27	84	4.1	192	2	BE320194
C 28	84	4.1	198	7	CO310903
C 29	82.5	4.0	197	9	CR399937
C 30	81	4.0	198	1	AV902435
C 31	80	3.9	187	6	CD135364
C 32	79.5	3.9	168	5	BQ569612
C 33	79	3.9	193	4	BG382358
C 34	78.5	3.8	196	8	BH197534
C 35	78.5	3.8	197	2	BF472123
C 36	78	3.8	168	4	BG009910
C 37	78	3.8	160	7	CO427973
C 38	77.5	3.8	169	7	T82938
C 39	76.5	3.7	178	7	CV297531
C 40	76	3.7	172	1	AI369463
C 41	76	3.7	187	8	BH853006
C 42	76	3.7	189	7	N55666
C 43	75.5	3.7	173	1	AL785118
C 44	75	3.7	163	7	W17747
C 45	75	3.7	198	1	AA049560

ALIGNMENTS

RESULT 1
BE669639/c
LOCUS
DEFINITION
7e16e07.x1 NCI CGAP In24 Homo sapiens CDNA clone IMAGE:3282660 3'
Similar to TR:080828 080828 HYPOTHETICAL 88.8 KD PROTEIN. ;, mRNA
Sequence.

ACCESSION
BE669639
VERSION
BE669639.1 GI:10030180
KEYWORDS
EST.

SOURCE
Homo sapiens (human)

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 (bases 1 to 199)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

AUTHORS
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapsb-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LIML, send email to:
info@image.liml.gov

JOURNAL
Trace considered overall poor quality
Seq primer: -40UP from Gibco
High quality sequence stop: 1.
Location/Qualifiers
1. 199

FEATURES
source

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: www-bio.lnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Insert Length: 519 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers

FEATURES

source

```
1. 160
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:912260"
/issue_type="kidney"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Kid1"
/note="Vector: pAMP10; mRNA made from invasive kidney tumor, cDNA made by oligo-dT priming. Non-directionally cloned. Size-selected on agarose gel, average insert size 600 bp. Reference: Krizman et al. (1996) Cancer Research 56:5380-5383."
```

ORIGIN

Alignment Scores:

Pred. No.:	7.91e-20	Length:	160
Score:	290.00	Matches:	51
Percent Similarity:	96.23%	Conservative:	0
Best Local Similarity:	96.23%	Mismatches:	2
Query Match:	14.17%	Indels:	0
DB:	1	Gaps:	0

US-10-077-111-13 (1-384) x AA501581 (1-160)

```
QY 47 HisSerProLeuIysPheHisThrTyrAlaValHisCysCysPheSerProSerGly 66
    |||||
DB 160 CATCTTCATGTAAGTTTCACTACCTATGCTGCCACATGCTGCTGTTATCCCTTCAGGA 101
    |||||
QY 67 HisIleuAlaSerCysSerThrAspGlyThrThrValLeuThrPantThrGluAnGly 86
    |||||
DB 100 CATATTTTGCATGCTGTTCAACAGATGATGACCACTGCTCATGTAATGTAATGTA 41
    |||||
QY 87 GluMetLeuAlaValMetGluGlnProSerGlySerPro 99
    |||||
DB 40 CAGATGCTGCAGATGAGAACACCTAGTGCAGCCCT 2
    |||||
```

RESULT 4

AI415375 199 bp mRNA linear EST 09-FEB-1999
LOCUS mc35f10.x1 Soares mouse p3NMF19.5 Mus musculus cDNA clone
DEFINITION IMAGE:350539.3 similar to TR:023121 023121 F10G19.3 PROTEIN. ;

mRNA sequence.

ACCESSION AI415375
VERSION AI415375.1 GI:4258879
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 199)

TITLE Maira, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Stepien, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
COMMENT The WashU-NCI Mouse EST Project 1999
Unpublished (1999)
Contact: Maira, M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LINL; contact the IMAGE Consortium (info@image.lnl.gov) for further information. This clone was previously sequenced on the 5' end only, this new data is from the 3' end

Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
High quality sequence stop: 1.
Location/Qualifiers

FEATURES

source

```
1. 199
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:350539"
/dev_stage="19.5 dpc total fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares mouse p3NMF19.5"
/note="Vector: p7T3D (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAGTGGAGCGCGCATTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified p7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M. Fatima Bonaldo. RNA was kindly provided by Dr. Minoru Ko (Wayne State University)."
```

ORIGIN

Alignment Scores:

Pred. No.:	1.41e-19	Length:	199
Score:	289.00	Matches:	55
Percent Similarity:	92.42%	Conservative:	6
Best Local Similarity:	83.33%	Mismatches:	5
Query Match:	14.12%	Indels:	0
DB:	1	Gaps:	0

US-10-077-111-13 (1-384) x AI415375 (1-199)

```
QY 312 IleProAspGluPheIleCysProlIleThrAspGluLeuMetLysAspProValIleAla 331
    |||||
DB 199 ATCCCTGACGACGTTTCATGTCACAAATPACAGAGAACTCATGTAAGACCCCGTCATCGCA 140
    |||||
QY 332 SerAspGlyThrSerIleGluIleGluAlaMetGluAsnThrIleSerLysLysArg 351
    |||||
DB 139 TCAGATGCTACTCTTCACGAGAGAGAAAGCATCAAGCTGATCCAAAGAAAGCGT 80
    |||||
QY 352 ThrSerProMetThrAsnLeuValLeuProSerAlaValLeuThrProAsnArgThrLeu 371
    |||||
DB 79 ACCAGCCCATGTCATTAATTTGGCTCTTCATTCACGTGACTGACCCCAACAGACACTG 20
    |||||
QY 372 LysMetAlaIleAsnArg 377
    |||||
DB 19 AAGATGGCCATCAACCA 2
    |||||
```

RESULT 5

BF653223/c 190 bp mRNA linear EST 25-APR-2001
LOCUS 276923 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION BF653223
VERSION BF653223.1 GI:11918292

KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus

REFERENCE AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos. 1 (bases 1 to 190)

Smith, T. F. L., Grose, W. M., Freking, B. A., Roberts, A. J., Stone, R. T., Casas, E., Wray, J. B., White, J., Cho, J., Fahnenkrug, S. C., Bennett, G. L., Heaton, M. P., Laegreid, W. W., Rohrer, G. A., Chitko-McKown, C. G., Pette, G., Holt, I., Karamycheva, S., Liang, F., Quackenbush, J. and Keel, J. W.

TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle

JOURNAL Genome Res. 11 (4), 626-630 (2001)

MEDLINE 21180013

PUBMED 11282978

COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred v0.960904.e. Vector identified by cross_match with the -minscore 18 and -mismatch 12 options.

PCR primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTATCCGACGACGACG
Plate: 68 row: H column: 24
Seq primer: ATTAGGTGACACTATG.

FEATURES
source
1. 190
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH108"
/clone_lib="MARC 3BOV"
/note="Vector: PCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from marrow, alveolar
macrophage, ovary, fetal semitendoneous muscle, and fetal
longissimus muscle."

ORIGIN

Alignment Scores:

Pred. No.:	1.68e-11	Length:	190
Score:	209.50	Matches:	43
Percent Similarity:	88.8%	Conservative:	5
Best Local Similarity:	79.63%	Mismatches:	4
Query Match:	10.23%	Indels:	2
DB:	2	Gaps:	1

US-10-077-111-13 (1-384) x BF653223 (1-190)

QY 333 AspGlyTyrSerTyrGluYsgIuAlaMetGluAsnTrpIleSerIlySlySargThr 352
:::|||||

Db 178 GAGGGCTACTCATATGAAAAGAACATGGAATTCATGACGAGAAAGAACGTACC 119

QY 353 SerPromeThrAsnLeuValLeuPro-SerAlaValLeuThrProAsnArgThrLeu 372
:::|||||

Db 118 AGTCCCATGACCAATCTGTTCTTCTTACAGTGCACCTTACCAATGAGACTGTGA 59

QY 372 sMeAlaIleAsnArgTrpLeu---GluThrHisGlnIly 384
:::|||||

Db 58 AATGCCCATGATGATGGCTAACTGACACATCAAAAA 19

RESULT 6
CA587431 138 bp mRNA linear EST 12-JAN-2004
LOCUS LBE12P34 cDNA from mouse aorta Mus musculus cDNA, mRNA sequence.
ACCESSION CA587431
VERSION CA587431.1 GI:40792674
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 138)
REFERENCE Borang,S., Andersson,T., Theiln,A., Odeberg,U. and Lundberg,U.
Vascular gene expression in atherosclerotic plaque prone regions
analysed by representational difference analysis
JOURNAL Unpublished (2002)
COMMENT Contact: Andersson Tove
Department of Biotechnology

KTH
Teknikringen 34, plan 6, 100 44 Stockholm, Sweden
Tel: +46 8 790 71 29
Fax: +46 8 245452
Email: tove@biotech.kth.se
Representations (amplified cDNA) from plaque prone regions
Seq primer: CTA TGA CCA TGA TTA CGC CAA G.

FEATURES
source
1. 138
/organism="Mus musculus"
/mol_type="mRNA"
/strain="APOE-/- and LDLR-/- on C57BL/6x 129 background"
/db_xref="taxon:10090"
/sex="male"
/dev_stage="8 weeks old"
/clone_lib="cDNA from mouse aorta"
/note="Organ: aorta; Site 1: DpnII; Site 2: DpnII; cDNA
was prepared from whole aorta divided in atherosclerotic
plaque prone regions (aortic arch and abdominal aorta
proximal part) and less plaque prone regions (descending
thoracic aorta and abdominal aorta distal part). cDNA was
fragmented with DpnII, linker ligated and amplified to
generated starting material for representational
difference analysis (RDA). The two cDNA pools were
subjected to iterative RDA subtraction and amplification
to enrich for gene fragments differentially expressed at
early stages of atherosclerosis."

ORIGIN

Alignment Scores:

Pred. No.:	4.26e-08	Length:	138
Score:	174.00	Matches:	33
Percent Similarity:	91.8%	Conservative:	1
Best Local Similarity:	89.1%	Mismatches:	3
Query Match:	8.50%	Indels:	0
DB:	6	Gaps:	0

US-10-077-111-13 (1-384) x CA587431 (1-138)

QY 348 LyelYelYArGThrSerPromeThrAsnLeuValLeuProSerAlaValLeuThrPro 367
|||||

Db 5 AAGAAAGACGTCAGAGCCCATGCAATTTGGCTCTTCCTTCACTGATGACGCCCA 64

QY 368 AsnArgThrLeuLySweAlaIleAsnArgTrpLeuGluThrHisGlnIly 384
|||||

Db 65 AACAGACACTGAAAGATGACCATCAACGATGGCTGAGACGACGAGAAG 115

RESULT 7
BB323974 191 bp mRNA linear EST 11-JUL-2000
LOCUS BB323974 RIKEN full-length enriched, 4 days neonate male adipose
DEFINITION Mus musculus cDNA clone B430102G11 3', mRNA sequence.
ACCESSION BB323974
VERSION BB323974.1 GI:9032288
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 191)
REFERENCE Kono,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T.,
Carninci,P., Hori,F., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,
Hirozane,T., Hori,T., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,
Izawa,M., Kadoya,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N.,
Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C.,
Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H.,
Okazaki,Y., Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K.,
Shibata,K., Shibata,Y., Shigemoto,Y., Shingawa,A., Shiraki,T.,
Sogabe,Y., Sugahara,Y., Suzuki,H., Suuki,H., Tagawa,A.,
Takahashi,F., Tomimaga,N., Toya,T., Tanoda,Y., Watanishi,A.,
Watanabe,S., Yamamura,T., Yamane,I., Yano,R., Yasunishi,A.,
Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,M., Muramatsu,M. and
Hayashizaki,Y.

TITLE
JOURNAL
COMMENT
RIKEN Mouse ESTs (Konno, H., et al.)
Unpublished (2000)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagasaki, S.,
Sasaki, N., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
Thermotabilization and thermostabilization of thermostable enzymes by
trehalose and its application for the synthesis of full length
cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh, M., Kikunai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,
Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M.,
Okazaki, Y., and Hayashizaki, Y.
Automated filtration-based high-throughput plasmid preparation
system. Genome Res. 9 (5), 463-470 (1999)
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
13-44 (1999)
Please visit our web site (http://genome.ritc.riken.go.jp) for
further details.
Location/Qualifiers
1. 191
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="B430102G11"
/sex="male"
/issue_type="adipose"
/dev_stage="4 days neonate"
/lab_host="DH10B"
/clone_id="RIKEN full-length enriched, 4 days neonate
male adipose"
/note="Site 1: Sali; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 10.0 and subtraction to Rot = 229.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5' GAGAGAGAGATCTCCGAGTTAATTAATTATCCCCCCCC
3']. cDNA was cleaved with XhoI and BamHI. Vector: a
modified Bluescript KS(+) after bulk excision from Lambda
FLC I."

ORIGIN
Alignment Scores:
Pred. No.: 1.57e-05 Length: 191
Score: 151.00 Matches: 29
Percent Similarity: 86.49% Conservative: 5
Best Local Similarity: 78.38% Mismatches: 3
Query Match: 7.38% Indels: 0
DB: 2 Gaps: 0
US-10-077-111-13 (1-384) x BB323974 (1-191)
Qy 348 LyelylslysaqrThSerProMetThAsnleuValleuProSerAlaValleuThPro 367
Db 3 AAGAGAGAGAGCTTACGAGCCCATACAGATTTCCTCTCTTATTGATGATACCCCC 62
Qy 368 AsnaArgThleuylsMetAlaIleAsnaArgTrrleuGlnThrhIsglnlyg 384
Db 63 AACAGCACACTGAAGATCGCTTCAACCGATGCTGAGAGCCACGAGAG 113

RESULT 8
A1622018/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
A1622018 144 bp mRNA linear EST 22-APR-1999
466031E09.x4 486 - leaf primordia cDNA library from Hake lab Zea
mays cDNA, mRNA sequence.
A1622018
A1622018.1 GI:4646943
EST.
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 144)
Malbot, V.
Maize ESTs from various cDNA libraries sequenced at Stanford
University
Unpublished (1999)
Contact: Malbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: valbot@stanford.edu
Plate: 466031 row: E column: 09.
Location/Qualifiers
1. 144
/organism="Zea mays"
/mol_type="mRNA"
/cultiivar="B73"
/db_xref="taxon:4577"
/tissue_type="leaf primordia"
/dev_stage="P7-P11 leaf"
/lab_host="E.coli XL1-Blue MFG"
/clone_id="486 - leaf primordia cDNA library from Hake
lab"
/note="Organ: shoot; Vector: Lambda zap; Hake lab cDNA
library."

ORIGIN
Alignment Scores:
Pred. No.: 2.91e-05 Length: 144
Score: 146.50 Matches: 27
Percent Similarity: 78.26% Conservative: 9
Best Local Similarity: 58.70% Mismatches: 9
Query Match: 7.16% Indels: 1
DB: 1 Gaps: 1
US-10-077-111-13 (1-384) x A1622018 (1-144)
Qy 313 ProAspGluIbheIleCySProlIerhArgGluIleuMetLySaPProValIleAlaSer 332
Db 137 CCGAGCCATTTCATCTCCGCCCATCTTCAAGAGATGATGTCATCATCTCGCGCTCA 78
Qy 333 AspGlyTyrSerYrGlnlysgIuAlaMetGluAsnTrpIleSerlySlylsyAqrThr 352
Db 77 GATGGCTACACCTTACGAGCGGCAAGGCCATTGAGCTGTGAGCAGACGAG--ACG 21
Qy 353 SerProMetThAsnleu 358
Db 20 TCCCCGATGACCAACTTG 3
RESULT 9
CC724603/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
CC724603 179 bp DNA linear GSS 23-JUN-2003
OCUEK62TH_ZM_0_7_1_5_KB Zea mays genomic clone ZMBMa0421103,
genomic survey sequence.
CC724603
CC724603.1 GI:32143536
GSS.
Zea mays

ORGANISM	Zea mays Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta, Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea. 1 (bases 1 to 179)
REFERENCE	WhiteLAW,C.A., Quackebush,J., Van Aken,S., Utecherback,T., Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
AUTHORS	Consortium for Maize Genomics Unpublished (2002)
TITLE	JOURNAL
COMMENT	Other_GSSs: OGUKE6ZRV Contact: Cathy WhiteLaw TIGR
FEATURES	9712 Medical Center Drive, Rockville, MD 20850, USA Tel: 301-838-5843 Fax: 301-838-0208 Email: whiteLaw@cigr.org Seq primer: TR Class: sheared ends.
source	location/Qualifiers 1..179 /organism="Zea mays" /mol_type="Genomic DNA" /strain="B73" /db_xref="taxon:4577" /clone="ZMZMBM0421L03" /clone_1lb="ZM_0.7_1.5_KB" /note="vector: pBCSK-; Site_1: HincII, 0.7-1.5 kb methylation filtered genomic DNA library"
ORIGIN	
Alignment Scores:	
Pred. No.:	0.00301 Length: 179
Score:	138.00 Matches: 26
Percent Similarity:	62.96% Conservative: 8
Best Local Similarity:	48.15% Mismatches: 20
Query Match:	6.74% Indels: 0
DB:	Gaps: 0
US-10-077-111-13 (1-384) x CC724603 (1-179)	
OY	331 AAlasrAAGATYTYSerTyrGluIueGuaImetGluAentTrpIleSerlyeLyAs 350 ::: ::: ::: ::: :::
Db	177 GCCGGGAGCGATTACGTTCAGAGGCCCAAGGCATCACGGGTGTTGCACAGCCGCC 118 ::: ::: ::: ::: :::
OY	351 ArgThrSerPrometThrAsnLeuValIeuProSerAlaValIeuThrProknaArgThr 370 ::: ::: ::: ::: :::
Db	117 GACACGTGCGCGCATGCCAATCTCAAGCTCGAGCAGCTCGAGCTCACGCCCGAACAGGGCG 58 ::: ::: ::: ::: :::
OY	371 LeuLYsMeChAIaIlAsnArgrTrpLeuglurThrhAglnLYs 384 ::: ::: ::: ::: :::
Db	57 CTCCGCTCGCGCATTCGAGTGCGGAGCAGCAGCAGCGCG 16 ::: ::: ::: ::: :::
RESULT 10	
AA912881/c	82 bp mRNA linear EST 26-AUG-1998
LOCUS	AA912881
DEFINITION	0127b05.e1 Soares_NFL_T.GBC_S1 Homo sapiens cDNA clone IMAGE:1524659.3, similar to TR:O23121 O23121 F10G19.3 PROTEIN. ;, mRNA sequence.
ACCESSION	AA912881
VERSION	AA912881.1 GI:3052273
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 82) NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap. National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997) Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	

This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.jlnl.gov) for further information.

Insert length: 509 Std Error: 0.00

Seq primer: -40m13 fwd. RT from Amersham

High quality sequence stop: 1.

Location/Qualifiers

1. 82

/organism="Homo sapiens"

/mol_type="RNA"

/db_xref="taxon:9606"

/clone="IMAGE:1524659"

/lab_host="DH10B"

/clone.lib="Soares NFL T GBC SI"

/note="Organ: pooled; Vector: pTR73D-Pac (Pharmacia) with a modified polylinker; Site_1; Not 1; Site_2; Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NBHL19W, testis NHT, and B-cell NCI CGAP GCBI) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of 1.M.A.G.E. Clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Alignment Scores:

Pred. No.: 0.000142 Length: 82

Score: 136.00 Matches: 26

Percent Simlarity: 96.30% Conservative: 0

Best Local Simlarity: 96.30% Mismatches: 1

Query Match: 6.64% Indels: 0

DB: 1 Gaps: 0

US-10-077-111-13 (1-384) x AA912881 (1-82)

QY 312 ILAEPRAAGSLNHHELCEYSPROJLEHRAAGGULEMELTYAAGAPPROVALILEALA 331

DB 82 ATTCTGATGATATTTATGTCTCAATACCTAGAGAACTTATGAAAGACCCGGTCATCGCA 23

QY 332 SERAPGSLTYISERTYRGLU 338

DB 22 TCAGTGGCTATTCATATGAA 2

RESULT 11

AC211342 184 bp DNA linear GSS 09-AUG-2000

LOCUS AC211342

DEFINITION Oryza sativa (japonica cultivar-group) DNA, clone:NFA028 0.703_1A, 3', flanking sequence of Tos17 insertion in rice strain NFA028, genomic survey sequence.

ACCESSION AC211342

VERSION AC211342.1 GI:32358532

KEYWORDS GSS.

ORGANISM Oryza sativa (japonica cultivar-group)

SOURCE Oryza sativa (japonica cultivar-group)

1 Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Euphorbiaceae; Oryzae; Oryza.

REFERENCE 1 Miyao,A., Tanaka,K., Murate,K., Sawaki,H., Takeda,S., Abe,K., Shinozuka,Y., Onosato,K. and Hirochika,H. Target Site Specificity of the Tos17 Retrotransposon Shows a Preference for Insertion within Genes and against Insertion in Retrotransposon-Rich Regions of the Genome Plant Cell 15 (8), 1771-1780 (2003)

JOURNAL MEDLINE 22779046

PUBMED 12897251

2 (bases 1 to 184)

REFERENCE Miyao,A., Sawaki,M. and Hirochika,H. Direct Submission Submitted (16-APR-2002) Akio Miyao, National Institute of Agrobiological Sciences, Molecular Genetics, 2-1-2, Kamondai,


```

Class: sheared ends.
FEATURES
    source
        location/Qualifiers
            1..145
                /organism="Zea mays"
                /mol_type="genomic DNA"
                /strain="B73"
                /db_xref="taxon:4577"
                /clone="ZMMBTA0600M07"
                /clone_id="ZM_0.6_1.0 kb"
                /note="Vector: PCR4-TOPO; Site.1: EcoRI; 0.6-1.0 kb high
                Cot selected genomic DNA library"
ORIGIN
Alignment Scores:
Pred. No.: 0.03 Length: 145
Score: 117.00 Matches: 20
Percent Similarity: 69.57% Conservative: 12
Best Local Similarity: 43.48% Mismatch: 14
Query Match: 5.72% Indels: 0
DB: 9 Gaps: 0
US-10-077-111-13 (1-384) x CG460443 (1-145)
Qy 312 lileProaAgLuPhaIleCySProlIethrAgiLueMetIyAsProValIleAa 331
|||||
143 ATCCGGCGGCACTTCGCGGCCGATCTGCGAGCACTCATCGGAGACCCGCTCACGGCG 84
|||||
Qy 332 SerAaPgIyTyrSerTgIuLyAgLuAmeGluAenTIpIleSerIyLyAaRg 351
|||||
Db 83 CCGGCGGGAATCATCGACGACCGGAGAGCATCGAGCGCTGCTGAGACAGCGGCGCGCC 24
|||||
Qy 352 ThrSerPrometThraSm 357
|||||
Db 23 ACCTGCCCGCTCACCCAC 6
|||||
RESULT 14
LOCUS AW032239 189 bp mRNA linear EST 18-MAY-2000
DEFINITION EST175693 tomato callus, TAMU Lycopersicon esculentum cDNA clone
CLOC35P22, mRNA sequence.
ACCESSION AW032239
VERSION AW032239.1 GI:5890995
KEYWORDS EST.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 189)
Alcala,J., Vredalov,J., White,R., Matern,A.L., Vision,T.,
Holt,I.E., Liang,F., Upson,J., Craven,M.B., Bowman,C.L., Ahn,S.,
Romling,C.M., Fraser,C.M., Martin,G.B., Tankley,S.D. and
Giovannoni,J.
Generation of ESTs from tomato callus tissue
Unpublished (1999)
COMMENT CCGI
Contact: CCGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.
FEATURES
    source
        location/Qualifiers
            1..189
                /organism="Lycopersicon esculentum"
                /mol_type="mRNA"
                /cultivar="TA96"
                /db_xref="taxon:4081"
                /clone="CLOC35P22"
                /tissue_type="callus"
                /dev_stage="25-40 days old"
                /lab_host="XLI-Blue MRF"
                /clone_id="tomato callus, TAMU"
                /note="Vector: pBluescript SK(-); Site.1: EcoRI; Site.2:

```

ALIGNMENT SCORES:	ORIGIN
Alignment Scores:	
Pred. No.:	
Score:	0.116
Percent Similarity:	113.00
Best Local Similarity:	63.83%
Query Match:	42.55%
	5.52%
	2
	Gaps: 0
US-10-077-111-13 (1-384) x AW032239 (1-189)	
Qy	300 Gtuteuakrgthrluysvallyserlseuserseclyleproabngiuphailecyspro 319
Db	39 GAAATTTCTCAAAATATACAAATGGAAGAAATTCAGATACCCCTTATTATTATGGCCG 98
Qy	320 lIethrargtluuewctlyaspprovalilealaserapgltyrserlyrgulys 339
Db	99 ATTCTCTAGAGATAAAGAAAGATCCCTGACGATCCATCACTGGATGACATATGATCG 158
Qy	340 GtualaMetGtuanTtpile 346
Db	159 GAAATCATCGAAGAAATGATTA 179
RESULT 15	
LOCUS	BB186455 177 bp mRNA linear EST 30-JUN-2000
DEFINITION	BB186455 RIKEN full-length cDNA clone A330033K06 3', mRNA sequence.
VERSION	BB186455
KEYWORDS	BB186455.1 GI:8847026
SOURCE	EST.
ORGANISM	Mus musculus (house mouse)
REFERENCE	Mus musculus
AUTHORS	Emmariyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 177)
	Komno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Iizawa,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M., Iizawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N., Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M., Matsuyama,T., Miki,R., Mituno,Y., Nakamura,M., Oda,H., Ozaki,Y., Ono,T., Owa,C., Saito,H., Saki,C., Sato,K., Shibata,K., Shibata,Y., Shigemoto,Y., Shingawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H., Tagawa,A., Takahashi,F., Tomimaga,N., Toyota,T., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yamanaoka,I., Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
	RIKEN Mouse ESTs (Komno,H., et al.)
	Unpublished (2000)
TITLE	Contact: Yoshihide Hayashizaki
JOURNAL	Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
COMMENT	The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan Tel.: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/ Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoaka,S., Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Thermosensitization and thermoactivation of thermolabile enzymes by trihaloethane and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998) Itoh,M., Kikuchi,N., Akiyama,J., Shibata,K., Iizawa,M., Kawai,J., Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.

Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
 Carinci, P. and Hayaehizaki, Y.
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
 Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

FEATURES

location/Qualifiers

```
1..177
  /organism="Mus musculus"
  /mol_type="mRNA"
  /db_xref="taxon:10090"
  /clone="A330033K06"
  /sex="male"
  /tissue_type="spinal cord"
  /dev_stage="adult"
  /lab_host="DH10B"
  /clone_id="RIKEN full-length enriched, adult male spinal cord"
  /note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAGATCCAGAGCTTTTCTTTTCTTTTCTTTN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCGAGTATTAATTATCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from lambda FLC I."
```

ORIGIN

Alignment Scores:

Pred. No.:	0.168	Length:	177
Score:	111.00	Matches:	20
Percent Similarity:	88.46%	Conservative:	3
Best Local Similarity:	76.92%	Mismatches:	3
Query Match:	5.42%	Indels:	0
DB:	2	Gaps:	0

US-10-077-111-13 (1-384) x BB186455 (1-177)

QY	358	LeuValLeuProSerAlaValLeuThrProAsnArgThrLeuAlaMetAlaIleAsnArg	377
DB	19	TTGCTTTTCCCTTCACTACCTCCCAACAGACACTGAGAGATGCCATCAACCGA	78
QY	378	TrpLeuGluThrHisGln	383
DB	79	TGCCTGGAGACCCACGAG	96

Search completed: February 5, 2005, 18:04:55
 Job time : 4207 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 5, 2005, 14:57:03 ; Search time 230 Seconds
(without alignments)
2731.871 Million cell updates/sec

Title: US-10-077-111-13

Perfect score: 2047
Sequence: 1 MWLHTLADHGDDVNCACF.....LTPNRLKAINRLETHOK 384

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delpop 6.0 , Delext 7.0	

Searched: 1202784 segs, 818138359 residues

Total number of hits satisfying chosen parameters: 1407054

Minimum DB seq length: 0
Maximum DB seq length: 200

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DBY=xih
-Q=/cgn2_1/USFTO.spool/US10077111/runat_04022005_121141_7210/app_query.fasta_1.583
-DB=Issued_Patents_NA -OPMT=fastap -SUFFIX=p2nszlm200.rn1 -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptio -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MATEL=200 -USER=US10077111@cgn 1.1.177@runat_04022005_121141_7210 -NCPU=6
-ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -IONCLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PTUS.COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	85	4.2	191	4	US-09-513-999C-21644
2	79.5	3.9	169	4	US-09-016-434-417
3	66	3.2	154	4	US-09-313-294A-1052
4	60.5	3.0	183	4	US-09-902-540-8231
5	57	2.8	93	1	US-08-190-802A-24
6	57	2.8	93	3	US-08-477-346-24
7	57	2.8	93	3	US-08-473-089-24
8	57	2.8	93	3	US-08-487-072A-24
9	56	2.7	99	1	US-08-190-802A-23
10	56	2.7	99	3	US-08-477-346-23
11	56	2.7	99	3	US-08-473-089-23
12	56	2.7	99	3	US-08-487-072A-23

C	13	55.5	2.7	164	4	US-09-513-999C-16123	Sequence 16123, A
	14	55.5	2.7	192	4	US-09-270-767-30620	Sequence 30620, A
	15	55	2.7	93	1	US-08-190-802A-22	Sequence 22, Appl
	16	55	2.7	93	3	US-08-477-346-22	Sequence 22, Appl
	17	55	2.7	93	3	US-08-473-089-22	Sequence 22, Appl
	18	55	2.7	93	3	US-08-487-072A-22	Sequence 22, Appl
	19	55	2.7	151	4	US-09-313-294A-1053	Sequence 1053, Ap
C	20	55	2.7	194	4	US-09-507-765-10	Sequence 10, Appl
	21	54	2.6	169	4	US-09-401-064-145	Sequence 145, App
	22	54	2.6	186	4	US-09-513-999C-36540	Sequence 36540, A
	23	54	2.6	186	4	US-09-902-540-9414	Sequence 9414, Ap
	24	53.5	2.6	160	4	US-09-513-999C-35455	Sequence 35455, A
C	25	53.5	2.6	168	4	US-09-902-540-4618	Sequence 4618, Ap
	26	53	2.6	183	1	US-07-872-678A-43	Sequence 43, Appl
C	27	52.5	2.6	159	4	US-09-270-767-3874	Sequence 3874, Ap
C	28	52.5	2.6	159	4	US-09-270-767-19156	Sequence 19156, A
C	29	52.5	2.6	169	4	US-09-513-999C-25155	Sequence 25155, A
C	30	52.5	2.6	159	4	US-09-513-999C-25165	Sequence 25165, A
	31	52.5	2.6	195	4	US-09-248-796A-8053	Sequence 8053, Ap
	32	52	2.5	99	4	US-09-270-767-26751	Sequence 26751, A
	33	52	2.5	180	4	US-09-513-999C-35461	Sequence 35461, A
	34	52	2.5	183	1	US-07-872-678A-43	Sequence 43, Appl
	35	52	2.5	192	4	US-09-248-796A-13145	Sequence 13145, A
	36	52	2.5	195	3	US-08-717-294-2	Sequence 2, Appl1
	37	52	2.5	186	1	US-08-324-243-2	Sequence 2, Appl1
	38	52	2.5	196	1	US-08-532-390-2	Sequence 2, Appl1
	39	52	2.5	196	5	PCT-US95-11511-2	Sequence 2, Appl1
C	40	51.5	2.5	143	4	US-09-513-999C-30418	Sequence 30418, A
C	41	51.5	2.5	176	4	US-10-044-359-15	Sequence 15, Appl
C	42	51.5	2.5	195	4	US-09-270-767-31404	Sequence 31404, A
	43	51	2.5	146	4	US-09-468-253B-2	Sequence 2, Appl1
	44	51	2.5	146	4	US-09-513-999C-14749	Sequence 14749, A
	45	51	2.5	187	4	US-09-313-294A-5562	Sequence 5562, Ap

ALIGNMENTS

RESULT 1
US-09-513-999C-21644
; Sequence 21644, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J. B.
; APPLICANT: Duclert, A.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SEQ ID NO 21644
; LENGTH: 191
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 136
; OTHER INFORMATION: w=a or t
US-09-513-999C-21644
Alignment Scores:
Pred. No.: 0.258
Score: 85.00
Percent Similarity: 50.00%
Best Local Similarity: 36.00%
Query Match: 4.15%
DB: 4
Gaps: 0
US-10-077-111-13 (1-384) x US-09-513-999C-21644 (1-191)

Oy 100 ValaAValaCyglInPheSerProADPserThrCysLeuAlaSerGlyIaaIaaLaapGly 119
 Db 42 GTTAACTGCTGTTCCATCTCTCCCAAGATGCAGATGCTGCTCTTGCAGCTTGAGAGAAAG 101
 Oy 120 ThrValValLeuTriPasnAlaGlnSerTyrLysLeuTyrArgCysGlySerValLysAsp 139
 Db 102 TCGGCTCTTCTTATGAGACATGAGGTCCTTACACCTMAATTGGAAAGCTAAGAGGCCATCAA 161
 Oy 140 GlySerLeuAlaAlaCysAlaPheSerPro 149
 Db 162 AGCAGTGTGTCTCTTGTGACATTCCTCCCCC 191

```

RESULT 2
US-09-016-434-417
Sequence 417, Application US/09016434
Parent No. 6500938
GENERAL INFORMATION:
APPLICANT: Janice Au-Young
APPLICANT: Jeffrey J. Sellhammer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
TITLE OF INVENTION: PATHWAY GENE EXPRESSION
NUMBER OF SEQUENCES: 1490
CORRESPONDENCE ADDRESSES:
ADDRESS: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016_434
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 417:
SEQUENCE CHARACTERISTICS:
LENGTH: 169 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: ENDCNOT01
CLONE: 2137838
US-09-016-434-417

Alignment Scores:
Pred. No.:          0.975           Length:      169
Score:              79.50           Matches:     19
Percent Similarity: 53.57%         Conservative: 11
Best Local Similarity: 33.93%       Mismatches:  25
Query Match:        3.88%           Indels:      1
DB:                  4              Gaps:        1

US-10-077-111-13 (1-384) x US-09-016-434-417 (1-169)

```

```

Db      CTCATACAGACGTTCTTACAGATGATGATCCCTTAAGATTCTGACCTCTTAGAAGGAAGCGTC 62
Qy      Leu1a1Val1Met1Gln1ProSer1Gly1SerPro1Val1Arg1Val1Cys1Gln1Phe1Ser1Pro1Asp 108
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db      ATCTATACACTTCAAGACATACGGGA---CCTGCTTTACTGTTTCATTCTTCAAAGGT 119
Qy      Ser1Thr1Cys1Leu1a1Ser1Gly1Ala1Ala1Phe1Gly1Thr1Val1Val1Leu1Tyr 124
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db      GGAGACCTATTTCATCAGAGAGTCAAGACACACAGGCTTATATATGG 167

```

```

RESULT 3
US-09-313-294A-1052
; Sequence 1052, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Laligudi, Raghunath V.
; APPLICANT: Ico, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OR INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ. ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 1052
; LENGTH: 154
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: miac feature
; OTHER INFORMATION: Incyte ID No. 6476212 700550309H1
; US-09-313-294A-1052

```

Alignment Scores:	
Pred. No.:	36.7
Score:	66.00
Percent Similarity:	60.00%
Best Local Similarity:	43.33%
Query Match:	3.42%
DB:	4
Length:	15
Matches:	13
Conservative:	5
Mismatches:	12
Indels:	
Gaps:	0

US-10-077-111-13 (1-384) x US-09-313-294A-1052 (1-154)

Qy 99 ProValArgValCysLeuInpHisSerProApSerThrCysLeuAlaAspArgValAlaAsp 118
Db 38 CCGTCGCTGCGCTGATTTTACGTAGCGGATTCGAAATCGCGCATCGATCCAGAT 97
Qy 119 GlyThrValLeuTrpAsnAlaGlnSer 128
Db 98 GGAAGATTAAAGCTTGGCGTATACGAAT 127

```

RESULT 4
US-09-902-540-8231/c
; Sequence 8231, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; RILE REFERENCE: 38-10(1584)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 8231
; LENGTH: 183
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-8231

```


Alignment Scores:

Pred. No.: 232 Length: 183
Score: 60.50 Matches: 17
Percent Similarity: 53.66% Conservative: 5
Best Local Similarity: 41.46% Mismatches: 10
Query Match: 2.96% Indels: 9
DB: 4 Gaps: 2

US-10-077-111-13 (1-384) x US-09-902-540-8231 (1-183)

QY 345 TYPILSESLYSLYSATGTHSerpMetThrsAnleuValleuProSerAlaVal 364

Db 173 TGGATGTCGTGATGAAGAACGACG-----CGAGTCGTCTTCATCGCGCGTC 123

QY 365 -----LeuThProAsnArgThreuleuysMetAlaIleAsnArgTyr 378

Db 122 CATCCGAGCCACCACCACTGTCCATACCGAGCGCGGTGATGCTGACTGCG 63

QY 379 Leu 379

Db 62 CTC 60

RESULT 5

US-08-190-802A-24
Sequence 24, Application US/08190802A
Patent No. 5519003

GENERAL INFORMATION:

APPLICANT: Mochly-Rosen, Daria

APPLICANT: Ron, Dorit

TITLE OF INVENTION: WD-40 - Derived Peptides and Uses

NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:

ADDRESS: Delinger & Associates

STREET: P.O. Box 60850

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306-0850

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/190.802A

FILING DATE: 01-FEB-1994

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Fabian, Gary R.

REGISTRATION NUMBER: 33,875

REFERENCE/DOCKET NUMBER: 8600-0139

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 324-0880

TELEFAX: (415) 324-0960

INFORMATION FOR SEQ ID NO: 24:

SEQUENCE CHARACTERISTICS:

LENGTH: 93 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: RACK1 rV DNA Sequence, Fig. 1A

US-08-190-802A-24

Alignment Scores:

Pred. No.: 191 Length: 93

Score: 57.00 Matches: 10

Percent Similarity: 65.00% Conservative: 3

Best Local Similarity: 50.00% Mismatches: 7

Query Match:

DB: 2.78% Indels: 0
US-10-077-111-13 (1-384) x US-08-190-802A-24 (1-93)

QY 106 SerProAspSerThrcysleuAlaSerGlyAlaAlaAspGlyThrValValleuTyrPasn 125

Db 34 TCTCCAGATGATCCCTGCTGCTTCTGAGCGAGGATGCGCAGGCTATGCTGTGGAT 93

RESULT 6

US-08-477-346-24
Sequence 24, Application US/08477346
Patent No. 6262023

GENERAL INFORMATION:

APPLICANT: Mochly-Rosen, Daria

APPLICANT: Ron, Dorit

TITLE OF INVENTION: WD-40 - Derived Peptides and Uses

NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:

ADDRESS: Morrison & Foerster

STREET: 2000 Pennsylvania Avenue, NW

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20006-1812

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/477,346

FILING DATE: 07-JUN-1995

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/487,072

FILING DATE: 07-JUN-1995

ATTORNEY/AGENT INFORMATION:

NAME: MURASHIGE, KATE H.

REGISTRATION NUMBER: 29,959

REFERENCE/DOCKET NUMBER: 2550-0025.20

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 887-1500

TELEFAX: (202) 887-0763

INFORMATION FOR SEQ ID NO: 24:

SEQUENCE CHARACTERISTICS:

LENGTH: 93 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: RACK1 rV DNA Sequence, Fig. 1A

US-08-477-346-24

Alignment Scores:

Pred. No.: 191 Length: 93

Score: 57.00 Matches: 10

Percent Similarity: 65.00% Conservative: 3

Best Local Similarity: 50.00% Mismatches: 7

Query Match: 2.78% Indels: 0

DB: 3 Gaps: 0

US-10-077-111-13 (1-384) x US-08-477-346-24 (1-93)

QY 106 SerProAspSerThrcysleuAlaSerGlyAlaAlaAspGlyThrValValleuTyrPasn 125

Db 34 TCTCCAGATGATCCCTGCTGCTTCTGAGCGAGGATGCGCAGGCTATGCTGTGGAT 93

RESULT 7

US-08-473-089-24
; Sequence 24, Application US/08473089
; Patent No. 6342368
; GENERAL INFORMATION:
; APPLICANT: Mochly-Rosen, Daria
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 2000 Pennsylvania Avenue, NW
; CITY: Washington
; STATE: DC
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,089
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 2550-0025.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 93 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: RACK1 rv DNA Sequence, Fig. 1A
; US-08-473-089-24

Alignment Scores:
Pred. No.: 191 Length: 93
Score: 57.00 Matches: 10
Percent Similarity: 65.00% Conservative: 3
Best Local Similarity: 50.00% Mismatches: 7
Query Match: 2.78% Indels: 0
DB: 3 Gaps: 0

US-10-077-111-13 (1-384) x US-08-473-089-24 (1-93)
QY 106 SerProApSerThrCysLeuAlaSerGlyAlaAlaAspGlyThrValValLeuTrpAsn 125
Db 34 TCTCCAGATGATCCTCTGCTTCTGAGGCAAGATGGCCAGGCTATGCTGTGGAT 93

RESULT 8
US-08-487-072A-24
; Sequence 24, Application US/08487072A
; Patent No. 6423684
; GENERAL INFORMATION:
; APPLICANT: Mochly-Rosen, Daria
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 2000 Pennsylvania Avenue, NW
; CITY: Washington

STATE: DC
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,072A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 2550-0025.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 93 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHEICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: RACK1 rv DNA Sequence, Fig. 1A
US-08-487-072A-24

Alignment Scores:
Pred. No.: 191 Length: 93
Score: 57.00 Matches: 10
Percent Similarity: 65.00% Conservative: 3
Best Local Similarity: 50.00% Mismatches: 7
Query Match: 2.78% Indels: 0
DB: 3 Gaps: 0

US-10-077-111-13 (1-384) x US-08-487-072A-24 (1-93)
QY 106 SerProApSerThrCysLeuAlaSerGlyAlaAlaAspGlyThrValValLeuTrpAsn 125
Db 34 TCTCCAGATGATCCTCTGCTTCTGAGGCAAGATGGCCAGGCTATGCTGTGGAT 93

RESULT 9
US-08-190-802A-23
; Sequence 23, Application US/08190802A
; Patent No. 5519003
; GENERAL INFORMATION:
; APPLICANT: Mochly-Rosen, Daria
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: P.O. Box 60850
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-0850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/190,802A
; FILING DATE: 01-FEB-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:

```

; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33, 875
; REFERENCE/DOCKET NUMBER: 8600-0139
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0960
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEITICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: RACK1 r1v DNA Sequence, Fig. 1A
; US-08-190-802A-23

Alignment Scores:
Pred. No.: 282 Length: 99
Score: 56.00 Matches: 13
Percent Similarity: 50.00% Conservative: 3
Best Local Similarity: 40.62% Mismatches: 14
Query Match: 2.74% Indels: 2
DB: 1 Gaps: 1

US-10-077-111-13 (1-384) x US-08-190-802A-23 (1-99)

Qy 53 HisThrTyAlaValAlaHisCysCysePheSerPro-----SergIyHisIleLeuAla 70
Db 4 CATTCAGATGGGTCTGTGTCGCGCTTCCCCGAACAGCAGCAACCCATCATCGTC 63

Qy 71 SerCysSerThrAepGlyThrThrValLeuTrpAsn 82
Db 64 TCCTCGGATGGGACCAAGCTGTCAGGTGTGGAT 99

RESULT 10
; Sequence 23, Application US/08477346
; Patent No. 6262023
; GENERAL INFORMATION:
; APPLICANT: Mochly-Rosen, Daria
; APPLICANT: Ron, Dorit
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 2000 Pennsylvania Avenue, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,346
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/487,072
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 2550-0025.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
```

```

; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEITICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: RACK1 r1v DNA Sequence, Fig. 1A
; US-08-477-346-23

Alignment Scores:
Pred. No.: 282 Length: 99
Score: 56.00 Matches: 13
Percent Similarity: 50.00% Conservative: 3
Best Local Similarity: 40.62% Mismatches: 14
Query Match: 2.74% Indels: 2
DB: 3 Gaps: 1

US-10-077-111-13 (1-384) x US-08-477-346-23 (1-99)

Qy 53 HisThrTyAlaValAlaHisCysCysePheSerPro-----SergIyHisIleLeuAla 70
Db 4 CATTCAGATGGGTCTGTGTCGCGCTTCCCCGAACAGCAGCAACCCATCATCGTC 63

Qy 71 SerCysSerThrAepGlyThrThrValLeuTrpAsn 82
Db 64 TCCTCGGATGGGACCAAGCTGTCAGGTGTGGAT 99

RESULT 11
; Sequence 23, Application US/08473089
; Patent No. 6342368
; GENERAL INFORMATION:
; APPLICANT: Mochly-Rosen, Daria
; APPLICANT: Ron, Dorit
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 2000 Pennsylvania Avenue, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,089
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 2550-0025.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-0763
; TELEFAX: (202) 887-0763
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEITICAL: NO
; ANTI-SENSE: NO
```

```

; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: RACK1 r1v DNA Sequence, Fig. 1A
US-08-473-089-23

Alignment Scores:
Pred. No.: 282 Length: 99
Score: 56.00 Matches: 13
Percent Similarity: 50.00% Conservative: 3
Best Local Similarity: 40.62% Mismatches: 14
Query Match: 2.74% Indels: 2
DB: 3 Gaps: 1

US-10-077-111-13 (1-384) x US-08-473-089-23 (1-99)

QY 53 HieThrTyAlaValHisCyScyPheSerPro-----SerGlyHisIleLeuAla 70
D 4 CATTGCAAGTGGGTGCTTGTGTGTCGGCTTCTCCCGAAGCAGCAACCTATCATGTC 63
QY 71 SerCySerThrApGlyThrValLeuTPaen 82
D 64 TCCTGGGATGGAGCAAGCTGCTCAAGTGTGGAAT 99

RESULT 12
US-08-487-072A-23
; Sequence 23, Application US/08487072A
; Patent No. 6423684
; GENERAL INFORMATION:
; APPLICANT: Mochly-Rosen, Daria
; APPLICANT: Ron, Dorit
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
; TITLE OF INVENTION: Theeotf
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morlison & Foerster
; STREET: 2000 Pennsylvania Avenue, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent'n Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,072A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 2550-0025.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: RACK1 r1v DNA Sequence, Fig. 1A
US-08-487-072A-23

Alignment Scores:
Pred. No.: 282 Length: 99
Score: 56.00 Matches: 13
Percent Similarity: 50.00% Conservative: 3
Best Local Similarity: 40.62% Mismatches: 14
```

```

Query Match: 2.74% Indels: 2
DB: 3 Gaps: 1

US-10-077-111-13 (1-384) x US-08-487-072A-23 (1-99)

QY 53 HieThrTyAlaValHisCyScyPheSerPro-----SerGlyHisIleLeuAla 70
D 4 CATTGCAAGTGGGTGCTTGTGTGTCGGCTTCTCCCGAAGCAGCAACCTATCATGTC 63
QY 71 SerCySerThrApGlyThrValLeuTPaen 82
D 64 TCCTGGGATGGAGCAAGCTGCTCAAGTGTGGAAT 99

RESULT 13
US-09-513-999C-16123/C
; Sequence 16123, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; PATENT REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 16123
; LENGTH: 164
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 12
; OTHER INFORMATION: y=c or c
US-09-513-999C-16123

Alignment Scores:
Pred. No.: 779 Length: 164
Score: 55.50 Matches: 16
Percent Similarity: 48.15% Conservative: 10
Best Local Similarity: 29.63% Mismatches: 19
Query Match: 2.71% Indels: 9
DB: 4 Gaps: 2

US-10-077-111-13 (1-384) x US-09-513-999C-16123 (1-164)

QY 12 GLVAspApValAsnCyScy-----AlaPheSerPheSerLeuLeuAlaThrCySer 29
D 154 GGAGACCAAGTCTCGCTGTCTATCAGCTGGAGTGCAGCAATCATCTACTGAC 95
QY 30 LeuAspLysThrIleArgLeuTySerLeuArgAspPheThrGluLeuProHisSerPro 49
D 94 CTGGAA-----CTCCAAAGCCCAAGGATGCCCCACCTCAGCT 56
QY 50 LeuLysPheHisThrTyAlaValHisCyScyPheSer 63
D 55 TCCAAAGTGTAACTATAGTGTGAGCTACCATGCTGACC 14

RESULT 14
US-09-270-767-30620
; Sequence 30620, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
```

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30620
; LENGTH: 192
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-30620

Alignment Scores:
Pred. No.: 1.02e+03      Length: 192
Score: 55.50            Matches: 14
Percent Similarity: 47.06%      Conservative: 10
Best Local Similarity: 27.45%    Mismatches: 26
Query Match: 2.71%             Indels: 1
DB: 4                     Gaps: 1

US-10-077-111-13 (1-384) x US-09-270-767-30620 (1-192)

QY 230 GluHieglneulysglnphethrgluaspitpsergluglvalserthrtrpneu 249
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 19 GAGCCCTGGCTCAGATCGAGGTGCATTAATTGACATCGACGACACCGATTGGCTG 78

QY 250 Cye---AlaGlnspleuLysAspleuValGlyIlehelysMetAsnAsnIleAspGly 268
    ::| ::| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 79 GCTCAGTCCGTTCCAGCTGCCGCAATACGTGATCTGTCAATATACACAAGTTACTGGC 138

QY 269 LysgluleuLeuAsnLeuThrLysgluserIeu 279
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 139 GCTGCCCTTCCAGATTGGCTGTGATTAATCTT 171

RESULT 15
US-08-190-802A-22
; Sequence 22, Application US/08190802A
; Patent No 5519003
; GENERAL INFORMATION:
; APPLICANT: Mochly-Rosen, Daria
; APPLICANT: Ron, Dorit
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
; TITLE OF INVENTION: Theoreof
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: P.O. Box 60850
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-0850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/190,802A
; FILING DATE: 01-FEB-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 8600-0139
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 93 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ANTI-SENSE: NO
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: RACK1 r111 DNA Sequence, Fig. 1A
```

```
US-08-190-802A-22

Alignment Scores:
Pred. No.: 335          Length: 93
Score: 55.00            Matches: 10
Percent Similarity: 66.67%      Conservative: 4
Best Local Similarity: 47.62%    Mismatches: 7
Query Match: 2.69%             Indels: 0
DB: 1                     Gaps: 0

US-10-077-111-13 (1-384) x US-08-190-802A-22 (1-93)

QY 105 PheSerProAspSerThrCysLeuAlaSerGlyAlaAlaAspGlyThrValLeuTrp 124
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 31 TTCTCTCTGACACCGGCGAGATTGTCTGTGGGTCCGAGACACAGCATTAATTATGG 90

QY 125 Asn 125
    |||
Db 91 AAT 93

Search completed: February 5, 2005, 18:08:50
Job time : 232 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 5, 2005, 04:18:44 ; Search time 763 Seconds

(without alignments)
9868.822 Million cell updates/sec

Title: US-10-077-111-12

Perfect score: 1272

Sequence: 1 gattcgcgttcctcactcgcg.....acaccaaagtaagaatcc 1272

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: N_Geneseq_16Dec04:*
2: geneeqn1980s:*
3: geneeqn1980s:*
4: geneeqn2000s:*
5: geneeqn2001s:*
6: geneeqn2002s:*
7: geneeqn2002s:*
8: geneeqn2003s:*
9: geneeqn2003s:*
10: geneeqn2003s:*
11: geneeqn2003s:*
12: geneeqn2004s:*
13: geneeqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1272	100.0	1272	6	AAD45075 Human RET
2	977.2	76.8	1553	6	AAD45071 Human RET
3	965.8	75.9	1818	6	AAD45070 Human RET
4	964.2	75.8	1817	10	ADB47507 Human CDN
5	964.2	75.8	1844	4	AA158876 Human pol
6	964.2	75.8	1844	5	ADQ90908 DNA encod
7	964.2	75.8	1844	9	ADB48858 Human hum
8	894	70.3	1773	8	ABZ24707 Human cel
9	890.2	70.0	1811	10	ADB63535 Human CDN
10	878.8	69.1	1996	5	ABV29028 Human pro
11	878.8	69.1	1996	5	ABV23190 Human pro
12	877.8	69.0	1908	6	AAD45076 Human RET
13	867.8	68.2	1826	4	AA160662 Human pol
14	791	62.2	1623	4	AA164914 Beta-rian
15	782	61.5	1686	5	ADM19385 Novel hum
16	654	51.4	838	5	ADM19628 Novel hum
17	478.6	37.6	1291	4	AA158364 Human GTP
18	464	36.5	1901	6	AAD45073 Mouse RET
19	390	30.7	668	13	ADQ54461 Novel can
20	377.2	29.7	630	6	AAD45072 Human RET

C 21	335	26.3	446	5	ABV05038 Human pro
C 22	331.6	26.1	366	4	AA136816 Probe #55
C 23	331.6	26.1	366	4	AAK30903 Human bon
C 24	331.6	26.1	366	4	AAK05311 Human bra
C 25	331.6	26.1	366	4	ABS30581 Human liv
C 26	331.6	26.1	366	6	ABS05651 Human gen
C 27	329.8	25.9	409	5	ABV14207 Human pro
C 28	328.8	25.8	441	5	ABV35301 Human pro
C 29	328.8	25.8	441	5	ABV44135 Human pro
C 30	297	23.3	297	4	AA150015 Probe #18
C 31	297	23.3	297	4	AAK44005 Human bon
C 32	297	23.3	297	4	AAK18116 Human bra
C 33	297	23.3	297	4	ABS43659 Human liv
C 34	297	23.3	297	6	ABS18238 Human gen
C 35	291.6	22.9	419	3	AAH30188 Human col
C 36	179.6	14.1	466	4	ABA58107 Human foe
C 37	179.6	14.1	466	4	AA137710 Probe #63
C 38	179.6	14.1	466	4	AAK31840 Human bon
C 39	179.6	14.1	466	4	AAK06182 Human bra
C 40	179.6	14.1	466	4	ABS31528 Human liv
C 41	179.6	14.1	466	6	ABS06599 Human gen
C 42	176.8	13.9	180	4	ABA70710 Human foe
C 43	176.8	13.9	180	4	AA150879 Probe #19
C 44	176.8	13.9	180	4	AAK44908 Human bon
C 45	176.8	13.9	180	4	AAK18959 Human bra

ALIGNMENTS

RESULT 1	
AAAD45075	
ID	AAAD45075 standard; DNA; 1272 BP.
AC	
XX	AAAD45075;
DT	
XX	27-DEC-2002 (first entry)
DE	
XX	Human RET16.2 splice variant DNA.
XX	
KW	Human; RET16; intracellular signal; inflammation-related disease; asthma;
KW	rheumatoid arthritis; psoriasis; multiple sclerosis; neuroprotective;
KW	transplant rejection; chronic obstructive pulmonary disease; TNF-alpha;
KW	inflammatory bowel disease; Crohn's disease; immunosuppressive; cancer;
KW	inacute respiratory distress syndrome; candida; ulcerative colitis;
KW	autoimmune disease; cystic fibrosis; gene therapy; cytostatic; neoplasm;
KW	cellular migration disorder; cell proliferation disorder; calcification;
KW	hyperinsulinaemia; diabetes type 2; systemic lupus erythematosus; tumour;
KW	cardiovascular disease; Wegener's granulomatosis; atherosclerosis;
KW	thalassaemia; vasotropic; gene; ds.
XX	
OS	Homo sapiens.
XX	
FH	
FT	Key
FT	CDS
FT	Location/Qualifiers
FT	111..1265
FT	/*tag= a
FT	/product= "Human RET16.2 splice variant protein"
XX	
XX	WO200266494-A2.
PD	29-AUG-2002.
XX	
PF	15-FEB-2002; 2002WO-US005162.
XX	
PR	16-FEB-2001; 2001US-0269366P.
PR	29-MAY-2001; 2001US-0294181P.
XX	
PA	(BRIM) BRISTOL-MYERS SQUIBB CO.
XX	
PI	Toddernud CG, Finger JN, Rillema J;
XX	
DR	WPI; 2002-682760/73.
DR	P-PSDB; AAE8167.

XX New human, mouse or rat RET16 genes and proteins, involved in
PT intracellular signaling cascade, useful for in gene therapy, particularly
PT for treating e.g. ischemia, cystic fibrosis, autoimmune disease, cancers,
PT tumors or neoplasms.

PS Claim 1; Page 166; 175pp; English.

The invention relates to human, mouse or rat RET16 genes and proteins, involved in intracellular signaling cascade. The RET16 protein or polynucleotide is useful for treating an inflammation-related disease or disorder, e.g. rheumatoid arthritis, juvenile arthritis, psoriasis, asthma, ischaemia-reperfusion, multiple sclerosis, rejection of organ or tissue transplants, chronic obstructive pulmonary disease, inflammatory bowel disease, Crohn's disease, ulcerative colitis, acute respiratory distress syndrome, systemic lupus erythematosus, autoimmune disease, cystic fibrosis, cancers, tumours or neoplasms. This inflammation-related disease or disorder also includes disorders associated with aberrant activation of the TNF- α pathway, disorders associated with aberrant cellular migration, proliferation, metastasis, juvenile idiopathic arthritis, haematogenous metastases of tumour cells, hyperinflammation, diabetes type 2, atherosclerosis, cardiovascular disease, colon cancer, tumour progression, Wegener's granulomatosis, stem cell transplantation complications, ischaemia-reperfusion injury, thalassemia, acute lung injury, graft rejection, ischaemic heart, coronary artery calcification or allergic inflammation. RET16 DNA is used in gene therapy. The present sequence is human RET16.2 splice variant DNA

Sequence 1272 BP; 349 A; 269 C; 290 G; 364 T; 0 U; 0 Other;

Query Match	100.0%	Score 1272;	DB 6;	Length 1272;
Best Local Similarity	100.0%	Pred. NO. 0;		
Matches 1272; Conservative	0;	Mismatches	0;	Gaps 0;

QY	1	GAATTCGGCTTTTACCTGCGCGGACAGTGAACCCGACCCGCGTGGGACCTTTGAAGCG	60
DB	1	GAATTCGGCTTTTACCTGCGCGGACAGTGAACCCGACCCGCGTGGGACCTTTGAAGCG	60
QY	61	GATCCCGCGCGCGCGCGCTCTGACAGCTGTTTTTTCTTCAATAAAGAAATGCTAAAC	120
DB	61	GATCCCGCGCGCGCGCGCTCTGACAGCTGTTTTTTCTTCAATAAAGAAATGCTAAAC	120
QY	121	TGATTCACACATTAGCTGATCATGTGACGATGTCAACTGCTGCTTCCTCTTTCC	180
DB	121	TGATTCACACATTAGCTGATCATGTGACGATGTCAACTGCTGCTTCCTCTTTCC	180
QY	181	TCCTGGCTACTTGGCTCCTTGGACAAAACAATGGCTGTACTGTTAAGTACTTTACTG	240
DB	181	TCCTGGCTACTTGGCTCCTTGGACAAAACAATGGCTGTACTGTTAAGTACTTTACTG	240
QY	241	AACTGCAACATTCCTCATGGAAGTTTCATCTATGCTGACCTGCTGCTGTTTTCTCC	300
DB	241	AACTGCAACATTCCTCATGGAAGTTTCATCTATGCTGACCTGCTGCTGTTTTCTCC	300
QY	301	CTTCAGACATATTTTGGCATCGTGTTCACAAGATGGTACCACTGTCTTATGGAATACG	360
DB	301	CTTCAGACATATTTTGGCATCGTGTTCACAAGATGGTACCACTGTCTTATGGAATACG	360
QY	361	AAAATGACAGATGCTGGCAGTATGGAACAGCTTACTAGTGGACCCCTGTAGAGGTTTTGCC	420
DB	361	AAAATGACAGATGCTGGCAGTATGGAACAGCTTACTAGTGGACCCCTGTAGAGGTTTTGCC	420
QY	421	AGTTTTCCCGACACTCAAGTGTTTGGCATCAGGGGACGTGATGGAACCTGTGGTTTTGT	480
DB	421	AGTTTTCCCGACACTCAAGTGTTTGGCATCAGGGGACGTGATGGAACCTGTGGTTTTGT	480
QY	481	GGAAATGACAGTCAATCAAAATTATATAGATGTGATGTTAAAGATGCTCTTGGCGG	540
DB	481	GGAAATGACAGTCAATCAAAATTATATAGATGTGATGTTAAAGATGCTCTTGGCGG	540
QY	541	CATGTGCAATTTTCTCTTAATGGAAGCTCTTTTGTCACTGGCTCTCATGTGTGATTTAA	600
DB	541	CATGTGCAATTTTCTCTTAATGGAAGCTCTTTTGTCACTGGCTCTCATGTGTGATTTAA	600

QY	601	CAAGTGTGGGATGATPAAATGAGTGTCTGCATAGTGAAGAAAAGCAGATGATCTTGGAAATTA	660
Db	601	CAGTGTGGGATGATPAAATGAGTGTCTGCATAGTGAAGAAAAGCAGATGATCTTGGAAATTA	660
QY	661	CGTGTGGGATTTTCTTCAACAGCGAGTTTCATGTGAGAGAACAGAGTCTTCAGTTTTC	720
Db	661	CTGTGTGGGATTTTCTTCAACAGCGAGTTTCATGTGAGAGAACAGAGTCTTCAGTTTTC	720
QY	721	GACTGGGCATCATGTGTGTCAGGATTTGCCAAGTCAAAATTTGGATTGTTCTTTTACCATA	780
Db	721	GACTGGGCATCATGTGTGTCAGGATTTGCCAAGTCAAAATTTGGATTGTTCTTTTACCATA	780
QY	781	TCTTAGCAGGCGCACAAGACATCAGCTGAAGCAATTTACCGAAGATTGGTCAGAGGAGG	840
Db	781	TCTTAGCAGGCGCACAAGACATCAGCTGAAGCAATTTACCGAAGATTGGTCAGAGGAGG	840
QY	841	TCGTCTGCAACATGGCTTTGTGCAACAAGATTTAAAGATCTGTGTGATTTTCAAGATGA	900
Db	841	TCGTCTCAACATGGCTTTGTGCAACAAGATTTAAAGATCTGTGTGATTTTCAAGATGA	900
QY	901	ATAACATTGATGAGAAAAGAACTGTTGGAATCTTACAAAAGAAAGTCTGGCTGATGATTTGA	960
Db	901	ATAACATTGATGAGAAAAGAACTGTTGGAATCTTACAAAAGAAAGTCTGGCTGATGATTTGA	960
QY	961	AAATTGAATCTCTAGACCTGGTAGTAAGTGTCTAGAGAAAATTTGAAGAGCTCAGACCA	1020
Db	961	AAATTGAATCTCTAGACCTGGTAGTAAGTGTCTAGAGAAAATTTGAAGAGCTCAGACCA	1020
QY	1021	AGGTTAAATCCCTTTCTTCAGGAATTCCTGTGAAATTTATATGTCCAAATACTAGAGAC	1080
Db	1021	AGGTTAAATCCCTTTCTTCAGGAATTCCTGTGAAATTTATATGTCCAAATACTAGAGAAC	1080
QY	1081	TTATGAAGAATCCGGTCACTGGCATCAGATGGGCTTATCATATGAAAAGAAAGCATGTGAA	1140
Db	1081	TTATGAAGAATCCGGTCACTGGCATCAGATGGGCTTATCATATGAAAAGAAAGCATGTGAAA	1140
QY	1141	ATTGATGAGCAAAAAGAAAGGTACAAATCCCATGACAAAATCTTGTTCTTCTTCACGGG	1200
Db	1141	ATTGATGAGCAAAAAGAAAGGTACAAATCCCATGACAAAATCTTGTTCTTCTTCACGGG	1200
QY	1201	TACTTACACCAATAGGACTCTGTGAAAATGGCCATCAATAGATGGCTGGAGACACACCAA	1260
Db	1201	TACTTACACCAATAGGACTCTGTGAAAATGGCCATCAATAGATGGCTGGAGACACACCAA	1260
QY	1261	AGTAAAGAAATTC 1272	
Db	1261	AGTAAAGAAATTC 1272	

RESULT 2	
AAD45071	
ID	AAD45071 standard; DNA; 1553 BP.

AC AAD45071;

DT 27-DEC-2002 (first entry)

Human RET16 open reading frame (ORF) DNA.

KM Human, KRT16; intracellular signal; inflammation-related disease; aschma;
KM Rheumatoid arthritis; psoriasis; multiple sclerosis; neuroprotective;
KM Transplant rejection; chronic obstructive pulmonary disease; TNF-alpha;
KM Inflammatory bowel disease; Crohn's disease; immunosuppressive; cancer;
KM Inactive respiratory distress syndrome; cardiac; ulcerative colitis;
KM Autoimmune disease; cystic fibrosis; gene therapy; cytostatic; neoplasm;
KM cellular migration disorder; cell proliferation disorder; calcification;
KM hyperinsulinaemia; diabetes type 2; systemic lupus erythematosus; tumour;
KM cardiovascular disease; Wegener's granulomatosis; atherosclerosis;
KM Chlasessema; vasotrophic; open reading frame; ORF; ds.

OS Homo sapiens.

XX

[illegible]

Db	241	AACGCCACATTTCCATTGAAGTTTCAACTTAATGCTGTCCACTGCTGCTTTCTCC	300
Qy	301	CTTCAGACATATTTTGGCATCGTGTTCACACAGATGGTACCACTGCTCTATGAAATAC	360
Db	301	CTTCAGACATATTTTGGCATCGTGTTCACACAGATGGTACCACTGCTCTATGAAATAC	360
Qy	361	AAATATGACAGATCTGGCAGTATGGAACACGCTAGTGGGACGCCCTGTAGAGGTTGCC	420
Db	361	AAATATGACAGATCTGGCAGTATGGAACACGCTAGTGGGACGCCCTGTAGAGGTTGCC	420
Qy	421	AGTTTCCCCAGACTCCACGCTGTGGCATCAGGGGACGCTGATGGAACCTGTGGTTTGT	480
Db	421	AGTTTCCCCAGACTCCACGCTGTGGCATCAGGGGACGCTGATGGAACCTGTGGTTTGT	480
Qy	481	GGAAATGACAGCTATACAAATTTATATAGATGTGTAGTGTAAAGATGGCTCTTGGCGG	540
Db	481	GGAAATGACAGCTATACAAATTTATATAGATGTGTAGTGTAAAGATGGCTCTTGGCGG	540
Qy	541	CATGTGCATTTTCTCTATAGGAAGCTTCCTTGTCACTGCTCCTCATGTGTGATTTAA	600
Db	541	CATGTGCATTTTCTCTATAGGAAGCTTCCTTGTCACTGCTCCTCATGTGTGATTTAA	600
Qy	601	CAGTGTGGATGATATAATGAGGTGTCCGCTAGTGAATAAAGCACATGATCTTGGAATTA	660
Db	601	CAGTGTGGATGATATAATGAGGTGTCCGCTAGTGAATAAAGCACATGATCTTGGAATTA	660
Qy	661	CTGTGTGGATTTTTCCTTCACAGCCAGTTTCTGTGAGGAACAAGTCTTCAGTTTTC	720
Db	661	CTGTGTGGATTTTTCCTTCACAGCCAGTTTCTGTGAGGAACAAGTCTTCAGTTTTC	720
Qy	721	GACTGGCATCATGTGTGAGGATTTGCCAAGTCAAAATTTGGATTTGTTTCTTTTACCATA	780
Db	721	GACTGGCATCATGTGTGAGGATTTGCCAAGTCAAAATTTGGATTTGTTTCTTTTACCATA	780
Qy	781	TCTT-----	784
Db	781	TCTTAAAGTTTGAATTAATAATATAAAGTACATGAGTGGGCACTGTGCTCTGTCTGG	840
Qy	785	-----	784
Db	841	CTTGTCCTTTTCCCGTATGGGACAGATGCTAGTCTCAGGTCAGTGGATTAAGTCTGCA	900
Qy	785	-----	784
Db	901	TAGTATGATTAATAATACAGAAATATCTTCAACATTTGACTCAGACACACAGTATG	960
Qy	785	-----	784
Db	961	TCACAACCTTGCTTTTGCACCTAATACCCTTTTACTTGTCTACTGTTCATGTGACAAA	1020
Qy	785	-----AGCAAGGGGACAGAAACATC	804
Db	1021	CAGTGAACATCTGGCAATTTGACCTCGAACAACCTTTGCCAAGCAAGGGGACAGAAACATC	1080
Qy	805	AGCTGAACCAATTTACCGAAGATTTGTTCAGAGAGTGTCTCAACAATGGCTTTGTGCAC	864
Db	1081	AGCTGAACCAATTTACCGAAGATTTGTTCAGAGAGTGTCTCAACAATGGCTTTGTGCAC	1140
Qy	865	AAGATTTAAAGATCTGTGTGTATTTTCAAGATGAATTAACATGATGGAAGAAGACTGT	924
Db	1141	AAGATTTAAAGATCTGTGTGTATTTTCAAGATGAATTAACATGATGGAAGAAGACTGT	1200
Qy	925	TGAATCTTACAAAAGAAAGTGTGCTGTATTTTGAATAATGANTCTTGAAGCTGGGTA	984
Db	1201	TGAATCTTACAAAAGAAAGTGTGCTGTATTTTGAATAATGANTCTTGAAGCTGGGTA	1260
Qy	985	GTAAGTGTCTGAGAAAATTTGAAGAGCTCAGAACCAAGTTTAAATCCCTTTCTCAGAA	1044
Db	1261	GTAAGTGTCTGAGAAAATTTGAAGAGCTCAGAACCAAGTTTAAATCCCTTTCTCAGAA	1320
Qy	1045	TTCTGTGATTAATTAATGTCCAACTAACTAGAGAACTTATGAAGAATCCGATCTGCTAT	1104

Db 1321 TTCCTGATGATTATTAATGTCACATTAACAGAGAACTTATGAAAGATCCGTCATGCGAT 1380
Cc 1105 CAGATGGCTATTCATATGAAGAGAGCAATGAGAAATGGATCAGCAAAAGAAACGTA 1164
Cc 1381 CAGATGGCTATTCATATGAAGAGAGCAATGAGAAATGGATCAGCAAAAGAAACGTA 1440
Cc 1165 CAGATGGCTATTCATATGAAGAGAGCAATGAGAAATGGATCAGCAAAAGAAACGTA 1224
Cc 1441 CAGATGGCTATTCATATGAAGAGAGCAATGAGAAATGGATCAGCAAAAGAAACGTA 1500
Cc 1225 AATGGCTATTCATATGAAGAGAGCAATGAGAAATGGATCAGCAAAAGAAACGTA 1268
Cc 1501 AATGGCTATTCATATGAAGAGAGCAATGAGAAATGGATCAGCAAAAGAAACGTA 1544
RESULT 3
AAD45070
ID AAD45070 standard; DNA; 1818 BP.
XX
AC AAD45070;
XX
DT 27-DEC-2002 (first entry)
XX
DE Human RET16 DNA.
XX
XX Human; RET16; intracellular signal; inflammation-related disease; asthma;
Km rheumatoid arthritis; psoriasis; multiple sclerosis; neuroprotection;
Km transplant rejection; chronic obstructive pulmonary disease; TNF-alpha;
Km inflammatory bowel disease; Crohn's disease; immunosuppressive; cancer;
Km inacute respiratory distress syndrome; cardiac; ulcerative colitis;
Km autoimmune disease; cystic fibrosis; gene therapy; cytoskeletal; neoplasm;
Km cellular migration disorder; cell proliferation disorder; calcification;
Km hyperinflammation; diabetes type 2; systemic lupus erythematosus; tumour;
Km cardiovascular disease; Wegener's granulomatosis; atherosclerosis;
Km thalassemia; vasotropic; gene; ds.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH 148..1578
FT CDS
FT /*tag= a
FT /product= "Human RET16 protein"
XX
XX MO200266494-A2.
XX
XX 29-AUG-2002.
XX
XX 15-FEB-2002; 2002MO-US005162.
XX
XX 16-FEB-2001; 2001US-0269366P.
XX
XX 29-MAY-2001; 2001US-0294181P.
XX
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
XX Pi Todderud CG, Finger JN, Rillema J;
XX
XX WPI: 2002-682760/73.
XX
XX P-PSDB; AAE28163.
XX
XX PT New human, mouse or rat RET16 genes and proteins, involved in
XX
XX PT intracellular signaling cascade, useful for in gene therapy, particularly
XX
XX PT for treating e.g. ischemia, cystic fibrosis, autoimmune disease, cancers,
XX
XX PT tumors or neoplasms.
XX
XX
XX Claim 1; Page 147-148; 175pp; English.
XX
XX The invention relates to human, mouse or rat RET16 genes and proteins,
Cc involved in intracellular signaling cascade. The RET16 protein or
Cc polynucleotide is useful for treating an inflammation-related disease or
Cc disorder, e.g. rheumatoid arthritis, juvenile arthritis, psoriasis, or
Cc asthma, ischaemia-reperfusion, multiple sclerosis, rejection of organ or
Cc tissue transplants, chronic obstructive pulmonary disease, inflammatory
Cc bowel disease, Crohn's disease, ulcerative colitis, inacute respiratory

Cc distress syndrome, systemic lupus erythematosus, autoimmune disease,
Cc cystic fibrosis, cancers, tumours or neoplasms. This inflammation-related
Cc disease or disorder also includes disorders associated with aberrant
Cc activation of the TNF-alpha pathway, disorders associated with aberrant
Cc cellular migration, proliferation, metastasis, juvenile idiopathic
Cc arthritis, haematogenous metastases of tumour cells, hyperinflammation,
Cc diabetes type 2, atherosclerosis, cardiovascular disease, colon cancer,
Cc tumour progression, Wegener's granulomatosis, stem cell transplantation
Cc complications, ischaemia-reperfusion injury, thalassemia, acute lung
Cc injury, graft rejection, ischaemic heart, coronary artery calcification
Cc or allergic inflammation. RET16 DNA is used in gene therapy. The present
Cc sequence is human RET16 DNA
XX
XX
XX Sequence 1818 BP; 523 A; 373 C; 382 G; 540 T; 0 U; 0 Other;
XX
XX
XX Query Match 75.9%; Score 965.8; DB 6; Length 1818;
XX Best Local Similarity 81.8%; Pred. No. 1.7e-284;
XX Matches 1253; Conservative 0; Mismatches 2; Indels 276; Gaps 1;
XX
Cc 12 TCACCTGCGCGGACGTCGACCCGCGCGGACCTTGAAGGCGGATCCGCGG 71
Cc 49 TCACCTGCGCGGACGTCGACCCGCGCGGACCTTGAAGGCGGATCCGCGG 108
Cc 72 CCCCCGCTCTGCGAGCTGTTTCTTCAATTAAGACATGCTGAATTCACACA 131
Cc 109 CCCCCGCTCTGCGAGCTGTTTCTTCAATTAAGACATGCTGAATTCACACA 168
Cc 132 TTAGCTGATCAGTGTGAGATGTCACCTGCTGCTTCCCTTCCCTTGGTACT 191
Cc 169 TTAGCTGATCAGTGTGAGATGTCACCTGCTGCTTCCCTTCCCTTGGTACT 228
Cc 192 TGCTCTTGTGAGCAAAACAAATTCGCTGTACTGCTTACGTAATTAAGTGCACAT 251
Cc 229 TGCTCTTGTGAGCAAAACAAATTCGCTGTACTGCTTACGTAATTAAGTGCACAT 288
Cc 252 TGCTCTTGTGAGCAAAACAAATTCGCTGTACTGCTTACGTAATTAAGTGCACAT 311
Cc 289 TGCTCTTGTGAGCAAAACAAATTCGCTGTACTGCTTACGTAATTAAGTGCACAT 348
Cc 312 ATTTTGGCATGCTGTGTCACAGATGTCACCTGCTGCTTCCCTTCCCTTCCGACAT 371
Cc 349 ATTTTGGCATGCTGTGTCACAGATGTCACCTGCTGCTTCCCTTCCCTTCCGACAT 408
Cc 372 ATGCTGGCAGTGTGAGCAAGCTGAGTGCAGCTGAGGCTTCCGATTTTCCCA 431
Cc 409 ATGCTGGCAGTGTGAGCAAGCTGAGTGCAGCTGAGGCTTCCGATTTTCCCA 468
Cc 432 GACTCCAGCTGTTGGCATCAGGGGAGCTGATGAACTGTGTTTGTGAATGACAG 491
Cc 469 GACTCCAGCTGTTGGCATCAGGGGAGCTGATGAACTGTGTTTGTGAATGACAG 528
Cc 492 TCATTAACAAATTAATGATGTGTGATGTTAAAGATGCTTCCCTTCCCTTCCGAT 551
Cc 529 TCATTAACAAATTAATGATGTGTGATGTTAAAGATGCTTCCCTTCCCTTCCGAT 588
Cc 552 TCCTCTTAATGAAAGCTTCTTGTGATGCTGCTGATGTTGATTAACAGTGTGAT 611
Cc 589 TCCTCTTAATGAAAGCTTCTTGTGATGCTGCTGATGTTGATTAACAGTGTGAT 648
Cc 612 GATTAACAAATGATGCTGATGTTGATTAACAGATGTTGATTAACAGTGTGAT 671
Cc 649 GATTAACAAATGATGCTGATGTTGATTAACAGATGTTGATTAACAGTGTGAT 708
Cc 672 TTTTCTTCAAGCAAGCTTCTGATGAGAAACAGGCTTCAATTTTGTGATGATCA 731
Cc 709 TTTTCTTCAAGCAAGCTTCTGATGAGAAACAGGCTTCAATTTTGTGATGATCA 768
Cc 732 TGGTGCAGGATGTCAGCAAGTGAATTTGATTTTCAACATATCTT----- 784
Cc 769 TGGTGCAGGATGTCAGCAAGTGAATTTGATTTTCAACATATCTTAGGTTT 828
Cc 785 ----- 784

Db	829	GAATTTAAATATTAAGAACACTGAGTGGGACACTGTGTCCTGTTCTTGCGCTTGTCCTTT	888
Qy	785	-----	784
Db	889	TCCCATGATGGGACGATCTAGTCTCAAGGTCAGTGATAAGTCTGTCAATGATATGAT	948
Qy	785	-----	784
Db	949	ACTAATACGTGAATATATCTTACACATTGACTGACACACAGGTATGTCAACCTTGT	1008
Qy	785	-----	784
Db	1009	GCTTTTGCACTAATACCCTTTTACTTGTCCTGTCATGTGACAAACAGTGAATC	1068
Qy	785	-----ACCAAGCGCAGACAGAACATCACTGAAGCA	815
Db	1069	TGGCAATTTGACCTGGAAACACTTTGCCAACAGGGGCACAGAACATCACTGAAGCA	1128
Qy	816	TTTACCGAAGATTGCTCAGAGAGTGCCTCACAATGGCTTTTGTGCACAAAGTTTAAA	875
Db	1129	TTTACCGAAGATTGCTCAGAGAGTGCCTCACAATGGCTTTTGTGCACAAAGTTTAAA	1188
Qy	876	GATCTGTGGTATTTTCAAGATGATTAACATTTGATGGAAAAAGAACTGTGAATCTTACA	935
Db	1189	GATCTGTGGTATTTTCAAGATGATTAACATTTGATGGAAAAAGAACTGTGAATCTTACA	1248
Qy	936	AAAGAAAGTCGCGTGATGATTTTGAAATTTGAATCTCTAGGACTCGTAGTAAAGTCGTG	995
Db	1249	AAAGAAAGTCGCGTGATGATTTTGAAATTTGAATCTCTAGGACTCGTAGTAAAGTCGTG	1308
Qy	996	AGAAAAATTGAAGAGCTCAGAACCAAGTTAAATCCCTTTCTTCAGAAATTCCTGATGA	1055
Db	1309	AGAAAAATTGAAGAGCTCAGAACCAAGTTAAATCCCTTTCTTCAGAAATTCCTGATGA	1368
Qy	1056	TTTATATGTCCAATTAACTAAGAACTTATGAAAGTCCGGTCATCGATCGATGAGTGCAT	1115
Db	1369	TTTATATGTCCAATTAACTAAGAACTTATGAAAGTCCGGTCATCGATCGATGAGTGCAT	1428
Qy	1116	TCATATGAAAAAGAAAGCAATGGAAAAATTGGATCAGCAAAAAGAAAGTACAGTCCCATG	1175
Db	1429	TCATATGAAAAAGAAAGCAATGGAAAAATTGGATCAGCAAAAAGAAAGTACAGTCCCATG	1488
Qy	1176	ACAAATCTTGTCTTCTTCAGCGGTACTTACACCAATPAGGACTCTGAATAATGGCATC	1235
Db	1489	ACAAATCTTGTCTTCTTCAGCGGTACTTACACCAATPAGGACTCTGAATAATGGCATC	1548
Qy	1236	AATAGATGGCTGGAGACACACCAAAAGTAAA	1266
Db	1549	AATAGATGGCTGGAGACACACCAAAAGTAAA	1579

XX	RESULT 4
ADBA47507	
ID	ADBA47507 standard; cDNA, 1817 BP.
XX	
AC	ADBA47507;
XX	
DT	04-DEC-2003 (first entry)
XX	
DE	Human cDNA upregulated in dendritic cells SEQ ID NO 207.
XX	
KW	ss; gene; human; dendritic cells; high throughput; cancer;
KW	infectious disease; autoimmune disease; allergy;
KW	grat versus host disease; vaccine enhancing; gene therapy
XX	
OS	Homo sapiens.
XX	
PN	US2003134283-A1.
XX	
PD	17-JUL-2003.
XX	
PF	03-OCT-2001; 2001US-00971392.
XX	

PR	03-Oct-2000; 2000US-0237652P.
XX	(PETE/) PETERSON D. P.
PA	(PEAR/) PEARSON C I.
PA	(COCK/) COCKS B G.
XX	
PI	Peterson DP, Pearson CI, Cocks BG;
XX	
DR	WPI; 2003-662509/62.
XX	
PT	New combination comprises cDNAs that are differentially expressed in dendritic cells useful for preparing a composition for diagnosing or treating cancer, infectious disease, autoimmunity, allergy or graft versus host disease.
PT	
XX	
PS	Claim 1; SEQ ID NO 207; 28pp; English.
XX	
CC	The invention relates to a combination comprising cDNAs that are differentially expressed in dendritic cells (DC). Also included is a high throughput method for detecting differential expression of one or more cDNAs in a sample containing nucleic acids. The combination is useful for preparing a composition for diagnosing, treating and monitoring the treatment of cancer, infectious disease, autoimmunity, allergy or graft versus host disease, or for enhancing a vaccine. The present sequence represents a human CDNA upregulated in dendritic cells. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?docID=20030134283.
CC	
XX	
SQ	Sequence 1817 BP; 519 A; 373 C; 383 G; 542 T; 0 U; 0 Other;
	Query Match 75.8%; Score 964.2; DB 10; Length 1817;
	Best Local Similarity 81.8%; Pred. No. 5,2e-284;
	Matches 1252; Conservative 0; Mismatches 3; Indels 276; Gaps 1;
OY	12 TCACCTGGCGGCGCACAGTGAACCCGCACCGCCGTGGGCACCTTGAAAGCGCATCCCCGCCG 71
DB	49 TCACCTGGCGGCGCACAGTGAACCCGCACCGCCGTGGGCACCTTGAAAGCGCATCCCCGCCG 108
OY	72 CCCCOCGCTCCGCAAGGCGCTTTTCCTTCAATAAAGAATGTAAGAACTATTACACA 131
DB	109 CCCCOCGCTCCGCAAGGCGCTTTTCCTTCAATAAAGAATGTAAGAACTATTACACA 168
OY	132 TTAGCTGATCATGSGTAGAGATGTAACACTGCTGSCCTTCTCCTTCCCTTGGACT 191
DB	169 TTAGCTGATCATGSGTAGAGATGTAACACTGCTGSCCTTCTCCTTCCCTTGGACT 228
OY	192 TGCTTCCTTGGACAAAACAATTGGCTGTACTCTGTTAGTGACTTTACTGAACTGCCACAT 251
DB	229 TGCTTCCTTGGACAAAACAATTGGCTGTACTCTGTTAGTGACTTTACTGAACTGCCACAT 288
OY	252 TCTCCATTGAAGTTTCACTAACCTATGCTGCTCAGCTGCTGCTTTCTCCGCTTCAGAGACAT 311
DB	289 TCTCCATTGAAGTTTCACTAACCTATGCTGCTCAGCTGCTGCTTTCTCCGCTTCAGAGACAT 348
OY	312 ATTTTGGCATGCTGTCCAACAGATGTAACACTGTCTTATGGAATACTGAAAAATGACAG 371
DB	349 ATTTTGGCATGCTGTCCAACAGATGTAACACTGTCTTATGGAATACTGAAAAATGACAG 408
OY	372 ATGCTGCGAGTATGAGAAACAGCTTAGTGGACGCCCTGTGAGGCTTTGCACTTTTCCCA 431
DB	409 ATGCTGCGAGTATGAGAAACAGCTTAGTGGACGCCCTGTGAGGCTTTGCACTTTTCCCA 468
OY	432 GACTCCACGTSSTTTGGCATCGAGGGGCGAGCTGATGGAACCTGNGTTTTGTGAAATGACAG 491
DB	469 GACTCCACGTSSTTTGGCATCGAGGGGCGAGCTGATGGAACCTGNGTTTTGTGAAATGACAG 528
OY	492 TCATACAAATATATAGATGTGTAGTATGTTAAAGATGGCTCTTGGCGGCATGTGCATTT 551
DB	529 TCATACAAATATATAGATGTGTAGTATGTTAAAGATGGCTCTTGGCGGCATGTGCATTT 588
OY	552 TCTCCTAATGGAAGCTTCTTTGTCACTGGCTCTTCATGTGTGATTTAACAGTGGCAT 611

```
Db      589 TCTCCTAATGAAAGCTTCTTGTCTCATGCGCTCCTCATGTGATGATTTAAACAGTGTGGAT 648
Qy      612 GATAAAATGAGGTGTCTGATAGTGAAGAAAGACATGATCTTGGAAATTAACCTGTGGAT 671
Db      649 GATTAATGAGGTGTCTGATAGTGAAGAAAGACATGATCTTGGAAATTAACCTGTGGAT 708
Qy      672 TTTTCTTACAGCCAGTTTCTGATGAGAAAGAGTCTTCAAGTTTTCATCTGGATCA 721
Db      709 TTTTCTTACAGCCAGTTTCTGATGAGAAAGAGTCTTCAAGTTTTCATCTGGATCA 768
Qy      732 TGTGTGAGGATGCGCAAGTCAAAATTTGATGTTCTTTACCCCATCTT----- 784
Db      769 TGTGTGAGGATGCGCAAGTCAAAATTTGATGTTCTTTTACCATATCTTAAGTTT 828
Qy      785 ----- 784
Db      829 GAATTAATAATAAATACAGTGAAGTGGGCACTGTCTCTGTCTGTCTGTCTT 888
Qy      785 ----- 784
Db      889 TCCCATGATGAGGAGATGATGATCTCAGGGTCAAGTGAATAGTCTCATATATATGAT 948
Qy      785 ----- 784
Db      949 ACTAATATCTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1008
Qy      785 ----- 784
Db      1009 GCTTTTGACACTAATACCTTTTACTTGTCTGATGCTTCAATGGAACAAAGTGAATC 1068
Qy      785 -----AGCAAGCGCACAGAACTCAAGTGAAGCA 815
Db      1069 TGGCAATTTGACCTGGAACACTTTGCGCAAGCAAGGACAGAACATCACTGATGAACAA 1128
Qy      816 TTTACCGAAGATTGCTGAGAGAGAGTGTCTCAACAGGCTTGTGTGACAAAGATTTAAA 875
Db      1129 TTTACCGAAGATTGCTGAGAGAGAGTGTCTCAACAGGCTTGTGTGACAAAGATTTAAA 1188
Qy      876 GATCTTGTGTGTATTTTCAAGATGAATTAATTAATTAATTAATTAATTAATTAATTA 935
Db      1189 GATCTTGTGTGTATTTTCAAGATGAATTAATTAATTAATTAATTAATTAATTAATTA 1248
Qy      936 AAAAGAAAGTGTGCTGATGATTTGAAATTTGAATCTTCAAGACTGCGTAAAGTCTG 995
Db      1249 AAAAGAAAGTGTGCTGATGATTTGAAATTTGAATCTTCAAGACTGCGTAAAGTCTG 1308
Qy      996 AGAAGAAATTTGAAGAGTCAAGAACCAAGTTAAATCCCTTCTTCAAGAAATTTCTGATGAA 1055
Db      1309 AGAAGAAATTTGAAGAGTCAAGAACCAAGTTAAATCCCTTCTTCAAGAAATTTCTGATGAA 1368
Qy      1056 TTTATATGTCCAAATTAATTAAGAACTTATGAAAGATCCGGTCAATCGATCAGATGCTAT 1115
Db      1369 TTTATATGTCCAAATTAATTAAGAACTTATGAAAGATCCGGTCAATCGATCAGATGCTAT 1428
Qy      1116 TCATATGAAAAAGAGCAATGAAAAATTGATCAGCAAAAAAGAAAGAAAGCTTACAGTCCATG 1175
Db      1429 TCATATGAAAAAGAGCAATGAAAAATTGATCAGCAAAAAAGAAAGCTTACAGTCCATG 1488
Qy      1176 ACAAAATGTGTCTTCTTCAAGGCTACTTAACAACCAATTAAGAACTGAAAAATGCGCATC 1235
Db      1489 ACAAAATGTGTCTTCTTCAAGGCTACTTAACAACCAATTAAGAACTGAAAAATGCGCATC 1548
Qy      1236 AATGATGCTGAGACACACCAAAAGTAAA 1266
Db      1549 AATGATGCTGAGACACACCAAAAGTAAA 1579
```

```
RESULT 5
AA158876
ID AA158876 standard; cDNA; 1844 BP.
AC XX
NC AA158876;
XX XX
```

```
DT      22-OCT-2001 (first entry)
XX
XX Human polynucleotide SEQ ID NO 1079.
DE
XX
KW Human; nocotropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukemia; ss.
OS
XX Homo sapiens.
XX
XX WO200153112-A1.
XX
XX 26-JUL-2001.
XX
XX 26-DEC-2000; 2000MO-US034263.
XX
XX 23-DEC-1999; 99US-00471275.
XX 21-JAN-2000; 2000US-00488725.
XX 25-APR-2000; 2000US-00552317.
XX 20-JUN-2000; 2000US-00598042.
XX 19-JUL-2000; 2000US-00620312.
XX 03-AUG-2000; 2000US-00653450.
XX 14-SEP-2000; 2000US-00662191.
XX 19-OCT-2000; 2000US-00693036.
XX 29-NOV-2000; 2000US-00727344.
XX
XX (HYSEQ-) HYSEQ INC.
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,
XX Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao Q;
XX Zhou P, Goodrich R, Drmanac RT;
XX WPI; 2001-442253/47.
XX
XX P-PSDB; AAM39720.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders such
XX as central nervous system injuries.
XX
XX Claim 1; SEQ ID NO 1079; 10078bp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and the
XX encoded polypeptides (AA38642-AA42213) with nocotropic,
XX immunosuppressant and cyostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localized neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilization of the activities such as: Immune system suppression,
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX assays for receptor activity, arthritis and inflammation, leukemias and
XX C.N.S disorders. Note: The sequence data for this patent did not form
XX part of the printed specification
XX
XX Sequence 1844 BP; 535 A; 380 C; 389 G; 540 T; 0 U; 0 Other;
SQ
XX
XX Query Match 75.8%; Score 964.2; DB 4; Length 1844;
XX Best Local Similarity 81.8%; Pred. No. 5.2e-284;
XX Matches 125; Conservative 0; Mismatches 3; Indels 276; Gaps 1;
```

```
Qy      12 TCACCTGCGGCGGACGATGACCGCCCGTGGGACCTTGAAGCGGATCCGCGCG 71
Db      65 TCACCTGCGGCGGACGATGACCGCCCGTGGGACCTTGAAGCGGATCCGCGCG 124
Qy      72 CCCCCTCTCTGAGGCTGTTTCTTCAAAATTAAGAACTGTAAGATTCACACA 131
Db      125 CCCCCTCTCTGAGGCTGTTTCTTCAAAATTAAGAACTGTAAGATTCACACA 184
```

[illegible]

Accession	Gene	Protein	Species	Length (aa)	Weight (kDa)	PI	Ref
D01234	ADQ99098	ADQ99098 standard; cDNA; 1844 BP.	Human	1844	18.44	5.12	[1]
D01235	ADQ99099	ADQ99099; 23-SEP-2004 (first entry)	Human	1844	18.44	5.12	[2]
D01236	ADQ99100	DNA encoding human GPCR-like protein seqid 768.	Human	768	7.68	5.12	[3]
D01237	ADQ99101	ophthalmological; immunomodulatory; cytoprotic; antiatherosclerotic; antidiabetic; GPCR-like protein; ophthalmic disorder; nephritic disorder; neurological disorder; immunological disorder; nephritic disorder; hormonal dysfunction; cancer; atherosclerosis; diabetes; molecular weight marker; food supplement; human; ss.	Human	768	7.68	5.12	[4]
D01238	ADQ99102	Homo sapiens.	Human	768	7.68	5.12	[5]
D01239	ADQ99103	US6569662-B1.	Human	768	7.68	5.12	[6]
D01240	ADQ99104	27-MAY-2003.	Human	768	7.68	5.12	[7]
D01241	ADQ99105	19-JUL-2000; 2000US-00620312.	Human	768	7.68	5.12	[8]
D01242	ADQ99106	21-JAN-2000; 2000US-00488725.	Human	768	7.68	5.12	[9]
D01243	ADQ99107	25-APR-2000; 2000US-00552317.	Human	768	7.68	5.12	[10]
D01244	ADQ99108	(HYSE-) HYSEQ INC.	Human	768	7.68	5.12	[11]
D01245	ADQ99109	Tang YT, Zhou P, Drmanac RT; WPI; 2001-442255/47.	Human	768	7.68	5.12	[12]
D01246	ADQ99110	New G-protein-coupled receptor-like polypeptides and polynucleotides. useful for treating diseases of ophthalmic, neurological, immunological and nephritic systems and hormonal dysfunction, cancer, atherosclerosis and diabetes.	Human	768	7.68	5.12	[13]
D01247	ADQ99111	Example 2; SEQ ID NO 768; 92pp; English.	Human	768	7.68	5.12	[14]

CC polypeptide in tissue. The polypeptides can also be used as molecular
CC weight markers and as a food supplement. This sequence represents a human
CC polynucleotide of the invention.

XX Sequence 1844 BP; 535 A; 380 C; 389 G; 540 T; 0 U; 0 Other;

Query Match 75.8%; Score 964.2; DB 5; Length 1844;
Best Local Similarity 81.8%; Pred. No. 5.2e-284;
Matches 1252; Conservative 0; Mismatches 3; Indels 276; Gaps 1;

```
QY 12 TCACCTGCGCGGACGTCAGCCGACCCGCGGACCTTGAAAGCGAGATCCCGCG 71
DB 65 TCACCTGCGCGGACGTCAGCCGACCCGCGGACCTTGAAAGCGAGATCCCGCG 124
QY 72 CCCCCGCTCTGACGAGCTGTTTCTTCAATAAAGAACATGTAAGTGAATTCACACA 131
DB 125 CCCCCGCTCTGACGAGCTGTTTCTTCAATAAAGAACATGTAAGTGAATTCACACA 184
QY 132 TTAGCTGATCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 191
DB 185 TTAGCTGATCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 244
QY 192 TGCTCTGTCGACAAACAAATTCGCTGTAATGCTGTAATGCTGTAATGCTGTAAT 251
DB 245 TGCTCTGTCGACAAACAAATTCGCTGTAATGCTGTAATGCTGTAATGCTGTAAT 304
QY 252 TCTCATGTAAGTTTCAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 311
DB 305 TCTCATGTAAGTTTCAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 364
QY 312 ATTTTGGCATGCTGTTCAACAGATGTAATGTAATGTAATGTAATGTAATGTAAT 371
DB 365 ATTTTGGCATGCTGTTCAACAGATGTAATGTAATGTAATGTAATGTAATGTAAT 424
QY 372 ATGCTGGACAGTATGTAAGCAGCTGAGGAGCCTGAGGAGCTGAGGAGCTGAGG 431
DB 425 ATGCTGGACAGTATGTAAGCAGCTGAGGAGCCTGAGGAGCTGAGGAGCTGAGG 484
QY 432 GACTCCACGTTGTTGGCATCAGGGGACGTAATGTAATGTAATGTAATGTAAT 491
DB 485 GACTCCACGTTGTTGGCATCAGGGGACGTAATGTAATGTAATGTAATGTAAT 544
QY 492 TCATTAACAAATTAATGATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 551
DB 545 TCATTAACAAATTAATGATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 604
QY 552 TCTCCTAATGAGAGCTTCTTGTCACTGCTCCTCATGAGTATTAACAGTGGGAT 611
DB 605 TCTCCTAATGAGAGCTTCTTGTCACTGCTCCTCATGAGTATTAACAGTGGGAT 664
QY 612 GATTAATGAGAGTGTGTCATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 671
DB 665 GATTAATGAGAGTGTGTCATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 724
QY 672 TTTTCTTCAACAGCAGTTTCTGATGTAATGTAATGTAATGTAATGTAATGTAAT 731
DB 725 TTTTCTTCAACAGCAGTTTCTGATGTAATGTAATGTAATGTAATGTAATGTAAT 784
QY 732 TGTGTCAGAGTATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 784
DB 785 TGTGTCAGAGTATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 844
QY 785 ----- 784
DB 845 GAATTAATAATTAAGTACACTGAGTGGGACCTGCTCTGCTGCTGCTGCTGCTT 904
QY 785 ----- 784
DB 905 TCCCATGATGGGAGATGCTAGTCTCAGGCTCAGTGATGATGCTGCTGCTGCTGCT 964
QY 785 ----- 784
DB 965 ACTAATACAGAAATATCTTACACACTGATGCTCAGCACACAGGATATGCTCAACT 1024
```

```
QY 785 ----- 784
DB 1025 GCTTTTGACCTAATACCTTTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1084
QY 785 ----- 784
DB 1085 TGCGAATTTGACCTGGAACACCTTTCAGACGACGACGACGACGACGACGACG 1144
QY 816 TTTTACCGAAGATTGTCAGAGAGAGTCTGCAACATGCTTTGTCAGACAGATTT 875
DB 1145 TTTTACCGAAGATTGTCAGAGAGAGTCTGCAACATGCTTTGTCAGACAGATTT 1204
QY 876 GATCTTGTGTTGTTTCTTCAAGATGTAATGTAATGTAATGTAATGTAATGTAAT 935
DB 1205 GATCTTGTGTTGTTTCTTCAAGATGTAATGTAATGTAATGTAATGTAATGTAAT 1264
QY 936 AAGAAAGTCTGCTGATGATTTGAAATGTAATGTAATGTAATGTAATGTAATGTAAT 995
DB 1265 AAGAAAGTCTGCTGATGATTTGAAATGTAATGTAATGTAATGTAATGTAATGTAAT 1324
QY 996 AGGAAATTTGAGAGCTCAGACCAAGTTAAATCCCTTTCTTCAAGAAATTCCTGATG 1055
DB 1325 AGGAAATTTGAGAGCTCAGACCAAGTTAAATCCCTTTCTTCAAGAAATTCCTGATG 1384
QY 1056 TTTATATGTCATTAATCTAGAGACTTATGAAAGATCCGCTCATCCGATCAGATGCTAT 1115
DB 1385 TTTATATGTCATTAATCTAGAGACTTATGAAAGATCCGCTCATCCGATCAGATGCTAT 1444
QY 1116 TCATTAAGAAAGAGACCAATGGAATTTGATGACCAAAAGAAAGTCAAGTCCCATG 1175
DB 1445 TCATTAAGAAAGAGACCAATGGAATTTGATGACCAAAAGAAAGTCAAGTCCCATG 1504
QY 1176 ACAATCTTGTCTTCTTCTTCAAGGCTACTTACCAAAATGAGACTGTAATGAGCATC 1235
DB 1505 ACAATCTTGTCTTCTTCTTCAAGGCTACTTACCAAAATGAGACTGTAATGAGCATC 1564
QY 1236 AATGATGCTGAGACACACCAAAAGTAA 1266
DB 1565 AATGATGCTGAGACACACCAAAAGTAA 1595
RESULT 7
ADB48858
ID ADB48858 standard; cDNA; 1844 BP.
XX
AC ADB48858;
XX
DT 04-DEC-2003 (first entry)
XX
DE Novel human cDNA SEQ ID NO 768.
XX
KW 88; cancer; neurodegenerative disease; human.
XX
OS Homo sapiens.
XX
PN US2003104529-A1.
XX
PD 05-JUN-2003.
XX
PF 04-JAN-2002; 2002US-00037270.
XX
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 19-JUL-2000; 2000US-00620312.
XX
PA (ZHOU/) ZHOU P.
PA (TANG/) TANG Y T.
PA (LIU/) LIU C.
PA (ASUNDI/) ASUNDI V.
PA (DRMANAC/) DRMANAC R T.
XX
PI Zhou P, Tang YT, Liu C, Asundi V, Drmanac RT;
```

XX WPI: 2003-678194/64.
XX
PT New polynucleotide, useful for treating diseases e.g., cancer or
PT neurodegenerative diseases.
XX
PS Claim 1; SEQ ID NO 768; 99bp; English.
XX
CC The invention relates to a polynucleotide comprising a sequence given in
CC the specification, or its mature protein-coding portion, or its
CC complement. The polynucleotide is useful for treating diseases e.g.,
CC cancer or neurodegenerative diseases and many others listed in the
CC specification. The present sequence represents a novel human cDNA. Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html?docID=20030104529.
XX
SQ Sequence 1844 BP; 535 A; 380 C; 389 G; 540 T; 0 U; 0 Other;
Query Match 75.8%; Score 964.2; DB 9; Length 1844;
Best Local Similarity 81.8%; Pred. No. 5.2e-284;
Matches 1253; Conservative 0; Mismatches 3; Indels 276; Gaps 1;
QY 12 TCACTTCGCGGCACTGACCCGACCGCCGCGGACCTTGAAGCGGATCCCGCG 71
DB 65 TCACCTCGGGGCACTGACCCGACCGCCGCGGACCTTGAAGCGGATCCCGCG 124
QY 72 CCCCCGCTCTGCAAGCTGTTTCTTCAATAAAGAAATGTAATCTGATTCACACA 131
DB 125 CCCCCGCTCTGCAAGCTGTTTCTTCAATAAAGAAATGTAATCTGATTCACACA 184
QY 132 TTAGCTATGATGAGTGAAGATGATGATGATGATGATGATGATGATGATGATGAT 191
DB 185 TTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 244
QY 192 TGCTCTTGGACAAACAAATTCGCTGATCTGATGATGATGATGATGATGATGAT 251
DB 245 TGCTCTTGGACAAACAAATTCGCTGATCTGATGATGATGATGATGATGATGAT 304
QY 252 TCTTCATGAAAGTTTCACTATGATGATGATGATGATGATGATGATGATGATGAT 311
DB 305 TCTTCATGAAAGTTTCACTATGATGATGATGATGATGATGATGATGATGATGAT 364
QY 312 ATTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 371
DB 365 ATTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 424
QY 372 ATGCTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 431
DB 425 ATGCTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 484
QY 432 GACTCCAGTGTGTTGGATGATGATGATGATGATGATGATGATGATGATGATGAT 491
DB 485 GACTCCAGTGTGTTGGATGATGATGATGATGATGATGATGATGATGATGATGAT 544
QY 492 TCATACAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 551
DB 545 TCATACAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 604
QY 552 TCTCTAATGAAAGCTTCTTGTGATGATGATGATGATGATGATGATGATGATGAT 611
DB 605 TCTCTAATGAAAGCTTCTTGTGATGATGATGATGATGATGATGATGATGATGAT 664
QY 612 GATTAATAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 671
DB 665 GATTAATAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 724
QY 672 TTTTCTTCAAGCAAGTTTCTGATGATGATGATGATGATGATGATGATGATGATGAT 731
DB 725 TTTTCTTCAAGCAAGTTTCTGATGATGATGATGATGATGATGATGATGATGATGAT 784
QY 732 TGTGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 784

DB 785 TGTGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 844
QY 785 ----- 784
DB 845 GAATTAATAATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 904
QY 785 ----- 784
DB 905 TCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 964
QY 785 ----- 784
DB 965 ACTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1024
QY 785 ----- 784
DB 1025 GCTTTGACCTTAATACCTTTTACTGATGATGATGATGATGATGATGATGAT 1084
QY 785 -----AGCAAGCGCACAGAACATCAGCTGAGCA 815
DB 1085 TGGCAATTTGACCTGAGAACATCTTGGCAAGAGACACAGAACATCAGCTGAGCA 1144
QY 816 TTACCGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 875
DB 1145 TTACCGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1204
QY 876 GATCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 935
DB 1205 GATCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1264
QY 936 AAGAAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 995
DB 1265 AAGAAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1324
QY 996 AGGAAATTTGAAGAGTCAAGACCAAGTTAAATCCCTTTCTTCAAGAAATTCGATGAA 1055
DB 1325 AGGAAATTTGAAGAGTCAAGACCAAGTTAAATCCCTTTCTTCAAGAAATTCGATGAA 1384
QY 1056 TTTATATGTCATTAATTAAGAAATTCGCTGATGATGATGATGATGATGATGAT 1115
DB 1385 TTTATATGTCATTAATTAAGAAATTCGCTGATGATGATGATGATGATGATGAT 1444
QY 1116 TCATATGAAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1175
DB 1445 TCATATGAAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1504
QY 1176 ACAATCTTGTCTTCTTCTTCAAGCGGTAATTACCAAAATGATGATGATGATGAT 1235
DB 1505 ACAATCTTGTCTTCTTCTTCAAGCGGTAATTACCAAAATGATGATGATGATGAT 1564
QY 1236 AATGATGCTGAGACACACCAAAAGTAA 1266
DB 1565 AATGATGCTGAGACACACCAAAAGTAA 1595
RESULT 8
AB224707
ID AB224707 standard; cDNA; 1773 BP.
XX
XX AB224707;
XX
XX 07-APR-2003 (first entry)
XX
XX Human cell growth, differentiation and death protein CGPD-19 cDNA.
DE
DE CGPD-19; cell growth; cell differentiation; cell death; human;
KW cytosolic; antiarteriosclerotic; hepatotropic; antiinflammatory;
KW antiproliferative; antianemic; ophthalmologic; auditory; anticonvulsant;
KW cerebroprotective; nootropic; neuroprotective; antiparkinsonian;
KW neuroleptic; tranquilizer; immunosuppressive; anti-HIV; antiallergic;
KW antineumatic; antichylous; antidiabetic; dermatological; nephroprotective;
KW antineumatic; antichylous; antidiabetic; dermatological; nephroprotective;
KW antibacterial; fungicide; antiparasitic; protozoicide; antihelminthic;

Db 785 CCAAGTCAAAATTTGATGTTCTTTACCCATATCTTAGGTTGAATTAATAATAA 844
785 ----- 784
Db 845 AAGTACACTGATGGGCACTGCTCTCTGTTCTGCTTTTCCCATGATGGGCA 904
785 ----- 784
Db 905 GATGCTAGTCTCAGGGTCACTGTGATAAGTCTCATAGTATGATAGTAATACTGAGAA 964
785 ----- 784
Db 965 TATACCTTCAACATGTAGCTGACACACAGGTATGTCAACATTGCTTTTGCACTAA 1024
785 ----- 784
Db 1025 TACCCCTTTTACTGCTACTGTTCAATGACAAAAGTGACATCTGGCAATTTGACCT 1084
785 -----AGCAAGGCGCACAGAACATCAGCTGAGCAATTTACCGAAGTTG 829
1085 GGAACACTTTGGCAAGCAAGGAGACACAGAACATCAGCTGAGCAATTTACCGAAGTTG 1144
830 GTGAGAGAGGTGCTCTCAACATGCTTTGTGACACAGATTTTAAAGATCTTGTGGTAT 889
1145 GTGAGAGAGAGGTGCTCTCAACATGCTTTGTGACACAGATTTTAAAGATCTTGTGGTAT 1204
890 TTTCAAGATGATTAACATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 949
1205 TTTCAAGATGATTAACATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1264
950 TGATGATTTGAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1009
1265 TGATGATTTGAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1324
1010 GCTCAGGACCAAGGTTAAATCCCTTTCTTCAGGAATTCCTGATGAATTTATGTCCAT 1069
1325 GCTCAGGACCAAGGTTAAATCCCTTTCTTCAGGAATTCCTGATGAATTTATGTCCAT 1384
1070 AACTAGAGAACTTATGAAATATCCGATCATCCGATCATGATGATGATGATGATGATGATGAT 1129
1385 AACTAGAGAACTTATGAAATATCCGATCATCCGATCATGATGATGATGATGATGATGATGAT 1444
1130 AGCAATGAAAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1189
1445 AGCAATGAAAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1504
1190 TCCTTACGCGGTACTTACACCAAAATGAGCTTGAAATGGCCATCAATAGATGGCTGGA 1249
1505 TCCTTACGCGGTACTTACACCAAAATGAGCTTGAAATGGCCATCAATAGATGGCTGGA 1564
QY 1250 GACACACCAAAAGTAAA 1266
1565 GACACACCAAAAGTAAA 1581
Db
RESULT 10
ID ABV29028 standard; cDNA; 1996 BP.
XX ABV29028;
AC
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 29019.
XX
XX Human prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KM pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX

PD 23-AUG-2001.
XX
XX 20-FEB-2001; 2001WO-US005171.
PF
XX
XX 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA
PI Schlegel R, Endege WO, Monahan JE;
XX
XX WPI; 2001-662795/76.
DR
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 6145-6146; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (1) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (1) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (1) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX Sequence 1996 BP; 615 A; 378 C; 447 G; 543 T; 0 U; 13 Other;
SQ
Query Match 69.1%; Score 878.8; DB 5; Length 1996;
Best Local Similarity 80.7%; Pred. No. 7.9e-258;
Matches 1166; Conservative 0; Mismatches 2; Indels 276; Gaps 1;
QY 99 CAAATTAAGAAACATGCTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 158
Db 4 CAAATTAAGAAACATGCTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 63
QY 159 TGCTTGCTCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 218
Db 64 TGCTTGCTCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 123
QY 219 TACTGTTAGTGAATTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 278
Db 124 TACTGTTAGTGAATTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 183
QY 279 GTCCACTGCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 338
Db 184 GTCCACTGCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 243
QY 339 ACCACTGCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 398
Db 244 ACCACTGCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 303
QY 399 GGCAGCCCTGTAGGGTTTGGCAAGTTTCCAGAGTCTCAGCGTTTGGCATAGGGGCA 458
Db 304 GGCAGCCCTGTAGGGTTTGGCAAGTTTCCAGAGTCTCAGCGTTTGGCATAGGGGCA 363
QY 459 GCTGATGAACTGTGTTTGTGAAATGACAGTCAATACAAATTAATATAGATGATGATGATGAT 518
Db 364 GCTGATGAACTGTGTTTGTGAAATGACAGTCAATACAAATTAATATAGATGATGATGATGAT 423
QY 519 GTTAAAGATGCTCTTGGCGGCGATGCAATTTTCTCTTAATGAAAGCTTCTTGTCACT 578
Db 424 GTTAAAGATGCTCTTGGCGGCGATGCAATTTTCTCTTAATGAAAGCTTCTTGTCACT 483

```
QY 579 GGCCTCATGTGGTGAATTTAAGATGGATGATAAATGAGATGTCATAGTGA 638
DB 484 GGCTCTCATGTGGTGAATTTAAGATGGATGATAAATGAGATGTCATAGTGA 543
QY 639 AAGCAGATGATCTTGAATTAAGTCTGCGATTTTCTTCAAGCCAGTTTCTGATGA 698
DB 544 AAGCAGATGATCTTGAATTAAGTCTGCGATTTTCTTCAAGCCAGTTTCTGATGA 603
QY 699 GAACAGATGATCTTGAATTAAGTCTGCGATTTTCTTCAAGCCAGTTTCTGATGA 758
DB 604 GAACAGATGATCTTGAATTAAGTCTGCGATTTTCTTCAAGCCAGTTTCTGATGA 663
QY 759 TGAATGTTCTTCTTACCATATCTT----- 784
DB 664 TGAATGTTCTTCTTACCATATCTTAGAGTTTGAATTAATTAATTAAGTCTGATG 723
QY 785 ----- 784
DB 724 GGGCAGTGTGCTCTGTTCTGCTTGTGCTTTTCCATGATGGCAGATGCTAGTCA 783
QY 785 ----- 784
DB 784 GGGCAGTGTGATGATGCTGCTGATGATGATGATGATGATGATGATGATGATG 843
QY 785 ----- 784
DB 844 TTGACTAGCAGCAGCAGATGATGATGATGATGATGATGATGATGATGATG 903
QY 785 ----- 784
DB 904 GCTACGTGTTCAATGACAAAGAGTGAACATCTGCGAATTTGACCTGGAACCTTGC 963
QY 785 --AGCAAGGGGCGACAGAAACATCAGCTGAGCAATTTACCGAAGTTGCTCAGAGAGTGC 842
DB 964 CAAGCAAGGGGCGACAGAAACATCAGCTGAGCAATTTACCGAAGTTGCTCAGAGAGAT 1023
QY 843 GTCTCAACATGAGCTTTGTCAGCAAGATTTAAGATCTTGTGCTTATTTCAAGATGAT 902
DB 1024 GTCTCAACATGAGCTTTGTCAGCAAGATTTAAGATCTTGTGCTTATTTCAAGATGAT 1083
QY 903 AACATGATGAGAAAGAACTGTTGAATCTTACAAAGAAAGTCTGCTGATGATTTGAA 962
DB 1084 AACATGATGAGAAAGAACTGTTGAATCTTACAAAGAAAGTCTGCTGATGATTTGAA 1143
QY 963 ATTGAATCTTGAAGCTGCGTGAATGAAGTCTGAGAAATGAGAGTCTGAGCAAG 1022
DB 1144 ATTGAATCTTGAAGCTGCGTGAATGAAGTCTGAGAAATGAGAGTCTGAGCAAG 1203
QY 1023 GTTAAATCCCTTTCTTGAAGATCTTGAAGATTTAATGCTTCAATTAAGTGAAGCTT 1082
DB 1204 GTTAAATCCCTTTCTTGAAGATCTTGAAGATTTAATGCTTCAATTAAGTGAAGCTT 1263
QY 1083 ATGAAGATCCGCTCATGATGATGATGATGATGATGATGATGATGATGATGAT 1142
DB 1264 ATGAAGATCCGCTCATGATGATGATGATGATGATGATGATGATGATGATGAT 1223
QY 1143 TGAATGAGCAAAAGAAAGTGAAGTCCATGCAATCTTGTCTTCTTCAAGCGTA 1202
DB 1324 TGAATGAGCAAAAGAAAGTGAAGTCCATGCAATCTTGTCTTCTTCAAGCGTA 1283
QY 1203 CTTCACCAAAATGAGATCTTGAAGATGAGATGAGATGAGATGAGATGAGATGAG 1262
DB 1384 CTTCACCAAAATGAGATCTTGAAGATGAGATGAGATGAGATGAGATGAGATGAG 1443
QY 1263 TAAA 1266
DB 1444 TAAA 1447
```

RESULT 11
ABV23190
ID ABV23190 standard; cDNA; 1996 BP.

```
XX AC ABV23190;
XX AC
DT 16-SEP-2002 (first entry)
XX DE Human prostate expression marker cDNA 23181.
XX DE
XX DE Human prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX DE pharmacogenomic marker; gene; ss.
XX OS Homo sapiens.
XX OS
XX PN W0200160860-A2.
XX PN
XX PD 23-AUG-2001.
XX PD
XX PF 20-FEB-2001; 2001WO-US005171.
XX PF
XX PR 17-FEB-2000; 2000US-0183319P.
XX PR 16-MAR-2000; 2000US-0189662P.
XX PR 25-MAY-2000; 2000US-0207454P.
XX PR 09-JUN-2000; 2000US-0211314P.
XX PR 18-JUL-2000; 2000US-0219007P.
XX PR 13-DEC-2000; 2000US-0255281P.
XX PR
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PA
XX PI Schlegel R, Endege WO, Monahan JE;
XX PI
XX DR WPI; 2001-662795/76.
XX DR
XX PT Novel isolated nucleic acid molecule associated with cancerous state of
XX PT prostate cells and correlating with presence of prostate cancer, useful
XX PT for detecting presence of prostate cancer, stage of prostate cancer.
XX PT
XX PT Claim 1; Page 4166-4167; 11750pp; English.
XX PT
XX XX
XX CC The invention relates to an isolated nucleic acid molecule (I) comprising
XX CC a nucleotide sequence given in Tables 1-9 (ABV0010-ABV62213) of the
XX CC specification or its complement. (I) is useful for: (a) assessing whether
XX CC a patient is afflicted with prostate cancer; (b) monitoring the
XX CC progression of prostate cancer in a patient; (c) assessing the efficacy
XX CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
XX CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX CC (e) selecting a composition for inhibiting prostate cancer in a patient;
XX CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
XX CC determining whether prostate cancer has metastasized in a patient; (h)
XX CC assessing the aggressiveness or indolence of prostate cancer in a patient
XX CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX CC
XX SQ Sequence 1996 BP; 615 A; 378 C; 447 G; 543 T; 0 U; 13 Other;
XX SQ
XX Query Match 69.1%; Score 878.8; DB 5; Length 1996;
XX Best Local Similarity 80.7%; Pred. No. 7.9e-258;
XX Matches 1166; Conservative 0; Mismatches 2; Indels 276; Gaps 1;
```

```
Db      244  ACCACGTCTCTATGAAATCTGAAAAATGACAGATGCTGGCAGTGTGAAACGCTTAGT 303
Qy      399  GGGAGCCCTGTGAGGGGTTTGGCCAGTTTCCCGAGCTCCAGCGTTTGGCATGAGGGGCA 458
Db      304  GGGAGCCCTGTGAGGGGTTTGGCCAGTTTCCCGAGCTCCAGCGTTTGGCATGAGGGGCA 363
Qy      459  GCTGATGGAATCTGTGTTTGTGGAATGCAAGTCAATACAAATTATATAGATGTGTAGT 518
Db      364  GCTGATGGAATCTGTGTTTGTGGAATGCAAGTCAATACAAATTATATAGATGTGTAGT 423
Qy      519  GTTAAAGATGGCTCTTGGCGGCAATGTCATTTTCTCTAATGGAAGCTTTCTTGTCAC 578
Db      424  GTTAAAGATGGCTCTTGGCGGCAATGTCATTTTCTCTAATGGAAGCTTTCTTGTCAC 483
Qy      579  GGGTCCCATGATGTGATTTTAAAGTGTGGATGATTAATATGAGTGTGCAATAGGAA 638
Db      484  GGGTCCCATGATGTGATTTTAAAGTGTGGATGATTAATATGAGTGTGCAATAGGAA 543
Qy      639  AAAGCAGATGATCTTGGAAATTACCTGCTGCAATTTTCTTCAAGCCAGTTTCTGATGA 698
Db      544  AAAGCAGATGATCTTGGAAATTACCTGCTGCAATTTTCTTCAAGCCAGTTTCTGATGA 603
Qy      699  GAAACAAGTCTTCAATTTTTCGATGTCATATGTCAGAGATTGCCAAGTCAAAATT 758
Db      604  GAAACAAGTCTTCAATTTTTCGATGTCATATGTCAGAGATTGCCAAGTCAAAATT 663
Qy      759  TGAATGTTCTTTTACCATCTT----- 784
Db      664  TGAATGTTCTTTTACCATCTTAAAGTTTGAATTAATAATTAATAAGTACATGAGT 723
Qy      785  ----- 784
Db      724  GGGCAGTGTGCTCTCTGTTCTGGCTTGTGCTTTTCCATGATGGCGAGATGCTAGTCA 783
Qy      785  ----- 784
Db      784  GGGTCAAGTGAATAGTCTGATATGATATGATATGATATGATATGATATGATATGATCA 843
Qy      785  ----- 784
Db      844  TTGACTCAGCAGACAGGATATGTCACAACTTGCTTTGCACTTAATACCTTTTAATT 903
Qy      785  ----- 784
Db      904  GCTACTGCTTCATGAGCAAAAAGTAAATCTGGCAATTTGACTGGAACAATTGCTGC 963
Qy      785  --AGCAAGGCGCAGAGACATCAGCTGAGCAATTTTACGAAAGTTGGTCAGAGAGATC 842
Db      964  CAAGCAAGGCGCAGAGACATCAGCTGAGCAATTTTACGAAAGTTGGTCAGAGAGAT 1023
Qy      843  GTCTCAACATGCTGTTTGTGCAAAAGATTTAAAGATCTTGTGGTATTTTCAAGATGAT 902
Db      1024  GTCTCAACATGCTGTTTGTGCAAAAGATTTAAAGATCTTGTGGTATTTTCAAGATGAT 1083
Qy      903  AACATGATGGAAGAAAGTGTGATCTTACAAAAGAAAGTGTGCTGATGATTTGAA 962
Db      1084  AACATGATGGAAGAAAGTGTGATCTTACAAAAGAAAGTGTGCTGATGATTTGAA 1143
Qy      963  ATTGAATCTTGAAGCTGCGTAAAGTGAAGTGAAGAAATTGAAGAGCTCAGAGCAAG 1022
Db      1144  ATTGAATCTTGAAGCTGCGTAAAGTGAAGTGAAGAAATTGAAGAGCTCAGAGCAAG 1203
Qy      1023  GTTAAATCCCTTTTTCAGAAATCCCTGATGAAATTTATATGTCATTAATGAGAACTT 1082
Db      1204  GTTAAATCCCTTTTTCAGAAATCCCTGATGAAATTTATATGTCATTAATGAGAACTT 1263
Qy      1083  ATGAAAGATCCGCTCATCGCATGAGATGGCTATTCATATGAAAGAGCAATGGAAT 1142
Db      1264  ATGAAAGATCCGCTCATCGCATGAGATGGCTATTCATATGAAAGAGCAATGGAAT 1323
Qy      1143  TGGATGAGCAAAAAGAAAGTAAAGTCCCATGACAAATCTTGTCTTCTTCCAGCGGTA 1202
```

```
Db      1324  TGGATCAGCAAAAAGAAAGCTTCAAGTCCCATGACAAATCTTGTCTTCTTCCAGCGGTA 1383
Qy      1203  CTTACACCAAAATAGACCTCTGAAAAATGGCCATCAATAGATGGCTGAGACACACCAAAAG 1262
Db      1384  CTTACACCAAAATAGACCTCTGAAAAATGGCCATCAATAGATGGCTGAGACACACCAAAAG 1443
Qy      1263  TAAA 1266
Db      1444  TAAA 1447

RESULT 12
AAD45076
ID      AAD45076 standard; DNA; 1908 BP.
XX
AC      AAD45076;
XX
DT      27-DEC-2002 (first entry)
XX
DE      Human RET16.3 splice variant DNA.
XX
KW      Human; RET16; intracellular signal; inflammation-related disease; asthma;
KW      rheumatoid arthritis; psoriasis; multiple sclerosis; neuroprotective;
KW      transplant rejection; chronic obstructive pulmonary disease; TNF-alpha;
KW      inflammatory bowel disease; Crohn's disease; immunosuppressive; cancer;
KW      acute respiratory distress syndrome; cardiac; ulcerative colitis;
KW      autoimmune disease; cystic fibrosis; gene therapy; cytostatic; neoplasm;
KW      cellular migration disorder; cell proliferation disorder; calcification;
KW      hyperinsulinaemia; diabetes type 2; systemic lupus erythematosus; tumour;
KW      cardiovascular disease; Wegener's granulomatosis; atherosclerosis;
KW      thalassemia; vasotropic; gene; de.
XX
OS      Homo sapiens.
XX
FH      Key      Location/Qualifiers
FT      CDS      136..1644
FT      /tag= a
FT      /product= "Human RET16.3 splice variant protein"
XX
PN      WO200266494-A2.
XX
PD      29-AUG-2002.
XX
PF      15-FEB-2002; 2002WO-US005162.
XX
PR      16-FEB-2001; 2001US-0269366P.
PR      29-MAY-2001; 2001US-0294181P.
XX
PA      (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
PI      Todderud CG, Finger JN, Rillema J;
XX
DR      WPI, 2002-682760/73.
XX
DR      P-PSDB; AAE28168.
XX
PT      New human, mouse or rat RET16 genes and proteins, involved in
PT      intracellular signaling cascade, useful for in gene therapy, particularly
PT      for treating e.g. ischemia, cystic fibrosis, autoimmune disease, cancers,
PT      tumors or neoplasms.
XX
PS      Claim 1; Page 168-169; 175pp; English.
XX
CC      The invention relates to human, mouse or rat RET16 genes and proteins,
CC      involved in intracellular signaling cascade. The RET16 protein or
CC      polynucleotide is useful for treating an inflammation-related disease or
CC      disorder, e.g. rheumatoid arthritis, juvenile arthritis, psoriasis,
CC      asthma, ischaemia-reperfusion, multiple sclerosis, rejection of organ or
CC      tissue transplants, chronic obstructive pulmonary disease, inflammatory
CC      bowel disease, Crohn's disease, ulcerative colitis, inacute respiratory
CC      distress syndrome, systemic lupus erythematosus, autoimmune disease,
CC      cystic fibrosis, cancers, tumours or neoplasms. This inflammation-related
CC      disease or disorder also includes disorders associated with aberrant
CC      activation of the TNF-alpha pathway, disorders associated with aberrant
```

cellular migration, proliferation, metastasis, juvenile idiopathic
 arthritis, haemogenous metastases of tumour cells, hyperinsulinemia,
 diabetes type 2, atherosclerosis, cardiovascular disease, colon cancer,
 tumour progression, Wegener's granulomatosis, stem cell transplantation
 complications, ischaemia-reperfusion injury, thalassemia, acute lung
 injury, graft rejection, ischaemic heart, coronary artery calcification
 or allergic inflammation. RET16 DNA is used in gene therapy. The present
 sequence is human RET16.3 splice variant DNA

XX Sequence 1908 BP; 550 A; 399 C; 405 G; 554 T; 0 U; 0 Other;

Query Match 69.0%; Score 877.8; DB 6; Length 1908;

Best Local Similarity 77.9%; Pred. No. 1.6e-257;

Matches 1253; Conservative 0; Mismatches 2; Indels 354; Gaps 2;

12 TCACCTGCGGCGGACGTGACCGCGCGCGGACCTTGAGGCGGATCCCGGCG 71
 37 TCACCTGCGGCGGACGTGACCGCGCGCGGACCTTGAGGCGGATCCCGGCG 96
 72 CCCCCGCTCCGCGGCGGCTTTTCTTCAATAAAGAACTGGTGAATGATTCACACA 131
 97 CCCCCGCTCCGCGGCGGCTTTTCTTCAATAAAGAACTGGTGAATGATTCACACA 156
 132 TTACCTGATCATGAGTGAAGATGTCACCTGCTGCTCTTCTTCCCTTTGGCTACT 191
 157 TTACCTGATCATGAGTGAAGATGTCACCTGCTGCTCTTCTTCCCTTTGGCTACT 216
 192 TGCTCTTGACCAAAACAAATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 251
 217 TGCTCTTGACCAAAACAAATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 276
 252 TCTTCATTTGAAGTTTCACTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 311
 277 TCTTCATTTGAAGTTTCACTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 336
 312 ATTTGGCATGCTGTTCAACAGATGTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 371
 337 ATTTGGCATGCTGTTCAACAGATGTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 396
 397 ATGCTGGCAGTGTGAAACAGCTTATGTCGACGCTGTCGAGGCTGTCGAGGCTTTCCCA 456
 432 GACTCCACGCTGTTGGCATCAGGGGCGACCTGATGGAATGCTGCTGCTGCTGCTGCT 491
 457 GACTCCACGCTGTTGGCATCAGGGGCGACCTGATGGAATGCTGCTGCTGCTGCTGCT 516
 492 TCATCAAAATTAATAGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTC 551
 517 TCATCAAAATTAATAGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTC 576
 552 TCTCTTAATGAGACCTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 611
 577 TCTCTTAATGAGACCTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 636
 612 GATTAATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 671
 637 GATTAATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 696
 672 TTTTCTTCAACGCGAGTTTCTGATGAGAAACAGGCTCTTCAATTTTTCAGTGGCATCA 731
 697 TTTTCTTCAACGCGAGTTTCTGATGAGAAACAGGCTCTTCAATTTTTCAGTGGCATCA 756
 732 TGTGGTCAAGATTCAGCAAGTCAAAATTTGATGTTGTTTACCATATCTT----- 784
 757 TGTGGTCAAGATTCAGCAAGTCAAAATTTGATGTTGTTTACCATATCTTGGTTT 816
 785 ----- 784
 817 GAATTAATAATAAAGTACAGTGGGCACTGTGCTCTGTTGGCTTGTGCTTTT 876
 785 ----- 784

877 TCCCATGATGAGGAGATGCTAGTCTCAGGCTCAGTGGATTAAGTCTGATAGTATATGAT 936
 785 ----- 784
 937 ACTAATAGTGAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 996
 785 ----- 784
 997 GCTTTTGACCTTAATACCTTTTACTTGTCTACTGTTTCAATGAGCAAAACAGTAACATC 1056
 785 ----- 784
 1057 TGGCAATTTGACCTTGAACACCTTTGCAAGCAAGGCGCACAGAACATCAGTGAAGCAA 1116
 816 TTTACCGAAGATTTGCTCAGAGAGGCTGCTCAACATGCTTTGTGCAACAAATTTTAA 875
 1117 TTTACCGAAGATTTGCTCAGAGAGGCTGCTCAACATGCTTTGTGCAACAAATTTTAA 1176
 876 GATCTGTTGCTATTTTCAAGATGAATTAATTAATTAATTAATTAATTAATTAATTAAT 935
 1177 GATCTGTTGCTATTTTCAAGATGAATTAATTAATTAATTAATTAATTAATTAATTAAT 1236
 936 AAGAAAGTCTGCTGATGATTTGAATTTG----- 966
 1237 AAGAAAGTCTGCTGATGATTTGAATTTG----- 1296
 967 ----- 977
 1297 ACTGACCTTCAACCTCTGAGGCTCAAGTATCTCTACCTGCGCTCAATCTTAAGAA 1356
 978 CTGCGTAGTAAGTCTGAGAAATTTGAAGCTCAGACCAAGGTTAAATCCCTTCT 1037
 1357 CTGCGTAGTAAGTCTGAGAAATTTGAAGCTCAGACCAAGGTTAAATCCCTTCT 1416
 1038 TCAGAAATTCCTGATGATTTATATGTCATTAATTAATTAATTAATTAATTAATTAAT 1097
 1417 TCAGAAATTCCTGATGATTTATATGTCATTAATTAATTAATTAATTAATTAATTAAT 1476
 1098 ATGCGATCAGATGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1157
 1477 ATGCGATCAGATGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1536
 1158 AAGCTAAGTCTGCTGATGATTTGAATTTG----- 1217
 1537 AAGCTAAGTCTGCTGATGATTTGAATTTG----- 1596
 1218 ACTCTGAATAATGCGCATCAATAGATGCTGAGACACCAAAAGTAA 1266
 1597 ACTCTGAATAATGCGCATCAATAGATGCTGAGACACCAAAAGTAA 1645

RESULT 13
 AA160662/c
 ID AA160662 standard; cDNA; 1826 BP.
 XX
 AA160662;
 XX
 22-OCT-2001 (first entry)
 XX
 Human polynucleotide SEQ ID NO 4651.
 DE
 Human, noctropin; immunosuppressant; cytotoxic; gene therapy; cancer;
 peripheral nervous system; neuropathy; central nervous system; CNS;
 Alzheimer's; Parkinson's disease; Huntington's disease; haemoblastic;
 amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemoblastic;
 chemoblastic; thrombolytic; drug screening; arthritis; inflammation;
 leukaemia; ss.
 XX
 XX Homo sapiens.
 OS
 XX
 XX W020015312-A1.
 XX
 XX 26-JUL-2001.

XX 26-DEC-2000; 2000MO-US034263.
PF
XX
PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX
XX (HYSE-) HYSEQ INC.
XX
PI Tang YT, Lau Z, Ahundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QH;
PI Zhou P, Goodrich R, Drmanac RT;
XX
DR MPI; 2001-442253/47.
DR P-PSDB; AAM41506.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX
XX Claim 1; SEQ ID NO 4651; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and the
CC encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Actin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening
CC assays for receptor activity, arthritis and inflammation, leukaemia and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
XX Sequence 1826 BP; 535 A; 382 C; 380 G; 529 T; 0 U; 0 Other;

Query Match 68.2%; Score 867.8; DB 4; Length 1826;
Best Local Similarity 79.6%; Pred. No. 1.8e-254;
Matches 1232; Conservative 0; Mismatches 27; Indels 288; Gaps 6;

QY 12 TCACCTGCGCGGACGTCACCGCCGCGGACCTTGAGGCGGATCCCGCGC 71
DB 1800 TCACCTGCGCGGACGTCACCGCCGCGGACCTTGAGGCGGATCCCGCGC 1741
QY 72 CCCCCGCTCTGCGAGCGTGTCTTCTTCAATAAAGACATGGTGAACGATTCACACA 131
DB 1740 CCCCCGCTCTGCGAGCGTGTCTTCTTCAATAAAGACATGGTGAACGATTCACACA 1681
QY 132 TTAGCTGATCATGTGTCAGATGTCACCTGTCGCTCTTCCCTTGGCTACT 131
DB 1680 TTAGCTGATCATGTGTCAGATGTCACCTGTCGCTCTTCCCTTGGCTACT 1621
QY 192 TGTCTCTTGACAAAACATTCGCTGTACTGTTAGTGAATTAAGTGCACAT 251
DB 1620 TGTCTCTTGACAAAACATTCGCTGTACTGTTAGTGAATTAAGTGCACAT 1561
QY 252 TCTTCATTGAAGTTTCACTACCTATGCTGCCACTGCTGCTGTCTCCCTTCAGACAT 311
DB 1560 TCTTCATTGAAGTTTCACTACCTATGCTGCCACTGCTGCTGTCTCCCTTCAGACAT 1501
QY 312 ATTTGGCATCGTGTCAACAGATGTCACCTGCTCTAGTGAATTAAGTGCACAG 371
DB 1500 ATTTGGCATCGTGTCAACAGATGTCACCTGCTCTAGTGAATTAAGTGCACAG 1441

QY 372 ATGCTGGCAGTATGGAACAGCCCTAGTGGCAGCCCTGTGAGGGTTTGCCAGTTTCCCA 431
DB 1440 ATGCTGGCAGTATGGAACAGCCCTAGTGGCAGCCCTGTGAGGGTTTGCCAGTTTCCCA 1381
QY 432 GACTTCACGTTTGGCATCAGGGGACGTCATGTAACGTGTGTTTGGTAATGACAG 491
DB 1380 GACTTCACGTTTGGCATCAGGGGACGTCATGTAACGTGTGTTTGGTAATGACAG 1321
QY 492 TCATACAAATTATATAGATGTGTATGTTAAAGATGCTCTTGGCGCATGTGATTT 551
DB 1320 TCATACAAATTATATAGATGTGTATGTTAAAGATGCTCTTGGCGCATGTGATTT 1261
QY 552 TCTCTTAATGGAAGCTTCTTTGTCACTGCTCTCATGTGTGATTTAACGTGGGAT 611
DB 1260 TCTCTTAATGGAAGCTTCTTTGTCACTGCTCTCATGTGTGATTTAACGTGGGAT 1201
QY 612 GATTAATAAGGTGTCTGATAGTGAATAAGCATGATCTTGGATTACCTGCTGCGAT 671
DB 1200 GATTAATAAGGTGTCTGATAGTGAATAAGCATGATCTTGGATTACCTGCTGCGAT 1141
QY 672 TTTTCTTCAAGCAGCTTCTGATGAGAACAAAGTCTTCACTTTTTCGACTGCGATCA 731
DB 1140 TTTTCTTCAAGCAGCTTCTGATGAGAACAAAGTCTTCACTTTTTCGACTGCGATCA 1081
QY 732 TGTGTCAAGATTGCCAAGTCAAAATTTGATTTCTTTTACCATACTT----- 784
DB 1080 TGTGTCAAGATTGCCAAGTCAAAATTTGATTTCTTTTACCATACTTAGGTTT 1021
QY 785 ----- 784
DB 1020 GAAATTAATAATAAAGTACACTGATGAGGACTGTCTCTGCTGTGCTGTGCTTTT 961
QY 785 ----- 784
DB 960 TCCCATGATGGGACGATGCTAGTCAAGGTCATGTGATTAAGTCTCATATATATGAT 901
QY 785 ----- 784
DB 900 ACTAATGATGAGATATATCTTACACATTGACTCAGACACAGATATGACAACTGT 841
QY 785 ----- 784
DB 840 GCTTTTGACCTAATACCTTTTACTTGTCTACTGTTCATATGACAAAACGTGAACATC 781
QY 785 -----AGCAAGCGCACAGACATGACGTGAAGCA 815
DB 780 TGGCAATTTGACCTGGAACACCTTTCGCAAGACAGACACAGATCAGCTGAAGCA 721
QY 816 TTTACCGAAGATTGTCAGAGAGTCTCTCAACATGCTTTTGTCACAAAGTTTAA 875
DB 720 TTTACCGAAGATTGTCAGAGAGTCTCTCAACATGCTTTTGTCACAAAGTTTAA 661
QY 876 GATCTGTGTGATTTTCAAGATGATTAACATTGATGGAAGAACTGTGTAATCTTACA 935
DB 660 GATCTGTGTGATTTTCAAGATGATTAACATTGATGGAAGAACTGTGTAATCTTACA 601
QY 936 AAAGAAAGTCTGCTGATGATTTGAAATTTGAAATCTCTAGACTGCGTGTGAAGTCTG 995
DB 600 AAAGAAAGTCTGCTGATGATTTGAAATTTGAAATCTCTAGACTGCGTGTGAAGTCTG 541
QY 996 AGGAAATTTGAAGGTCAGAACCAAGTTAATCCCTTTTCAGAAATTCCTGATGA 1055
DB 540 AGGAAATTTGAAGGTCAGAACCAAGTTAATCCCTTTTCAGAAATTCCTGATGA 481
QY 1056 TTTATATGTCATTAATCTAGAGAACTTATGAAGATCCGCTCATGCGATCAAGTGGCTAT 1115
DB 480 TTTATATGTCATTAATCTAGAGAACTTATGAAGATCCGCTCATGCGATCAAGTGGCTAT 421
QY 1116 TCATATGAAGAAGAACATGGAATTTGG--ATCAGCAAAAAGAA--GATACAGTCC 1172
DB 420 TCATATGAAGAAGAACATGGAATTTGGATTCAGCAAAAAGAA--GATACAGTCC 361
QY 1173 ATGACAAATCTGTCTCTCTT-----CAGCGGTAC--TTACACCAATAGGACTCTGA 1224

Db 360 CCATGACAAATCCTGTCTTCCCTTCAGCGGTAACCTTAACCAATAGACTGTA 301
QY 1225 AAAT-GGGCATCATAGATGGCTGAGACACACCAAAAGTAAAGAT 1270
Db 300 AAAATGGCCATCAATAGATGGCTGAGACACCCCAAAAGTAAAT 254

RESULT 14
AA164914
ID AA164914 standard; cDNA; 1623 BP.
XX
XX AA164914;
AC
XX 04-DEC-2001 (first entry)
DT
XX Beta-transducin 41 coding sequence.
DE
XX Beta-transducin 41; cancer; HIV infection; gene therapy; ss.
KW
XX
XX Unidentified.
OS
XX
XX CN1300734-A.
FN
XX 27-JUN-2001.
PD
XX 21-DEC-1999; 99CN-00124285.
PF
XX 21-DEC-1999; 99CN-00124285.
PR
XX 21-DEC-1999; 99CN-00124285.
XX
XX (UYFU-) UNIV FUDAN.
PA
XX
XX Mao Y, Xie Y;
PI
XX WPI; 2001-537038/60.
DR
XX P-PSDB; AAG78660.
XX
XX New polypeptide-beta-transducin 41 for treating e.g. cancer and HIV
PT infection.
XX
XX
XX Claim 6; Page 24-25 (Disclosure); 32pp; Chinese.
PS
XX
XX The present invention provides the protein and coding sequences of beta-
CC transducin 41. The sequences can be used in the treatment of cancer and
CC HIV infection. The present sequence is the coding sequence of the
CC invention
XX
SQ Sequence 1623 BP; 485 A; 302 C; 335 G; 501 T; 0 U; 0 Other;

Query Match 62.2%; Score 791; DB 4; Length 1623;
Best Local Similarity 79.6%; Pred. No. 5.6e-231;
Matches 1102; Conservative 0; Mismatches 5; Indels 278; Gaps 3;

QY 160 GCTGTGCTCTCTCTCTTCCCTTGGCTACTGCTCTTGACAAAACAATTCGCTGT 219
Db 1 GCTGTGCTCTCTCTTCTCTTCCCTTGGCTACTGCTCTTGACAAAACAATTCGCTGT 60
QY 220 ACTGTAGCTGACTTACTGAACTGCCAATTCCTCAATGAACTTTCACTATGCTG 279
Db 61 ACTGTAGCTGACTTACTGAACTGCCAATTCCTCAATGAACTTTCACTATGCTG 120
QY 280 TCCACTGCTGCTGTTTCTCCCTTCAGACATATTTTGGCATGCTGTTCAACAGATGTA 339
Db 121 TCCACTGCTGCTGTTTCTCCCTTCAGACATATTTTGGCATGCTGTTCAACAGATGTA 180
QY 340 CCACTGTCTATGATAATAAGTGAAGATGCTGAGATGAGTGAACAGCTATGCTG 399
Db 181 CCACTGTCTATGATAATAAGTGAAGATGCTGAGATGAGTGAACAGCTATGCTG 240
QY 400 GCAGCCCTGTGAGGGTTTGGCCAGTTTTCGCCAGACTCCACGCTGTTGGCATCAGGGCAG 459
Db 241 GCAGCCCTGTGAGGGTTTGGCCAGTTTTCGCCAGACTCCACGCTGTTGGCATCAGGGCAG 300

QY 460 CTGATGAACTGTGCTTTTGTGGAATGACAGCTATCAATTAATATAGATGTAGT 519
Db 301 CTGATGAACTGTGCTTTTGTGGAATGACAGCTATCAATTAATATAGATGTAGT 360
QY 520 TTAAGATGAGCTCTTGGCGCATGTGCAATTTTCTCTTAATGAAAGCTTCTTTGTCACTG 579
Db 361 TTAAGATGAGCTCTTGGCGCATGTGCAATTTTCTCTTAATGAAAGCTTCTTTGTCACTG 420
QY 580 GCTTCTCATGTGTGATTTTAACAGTGTGGATGATAAAGAGTGTCTCATATGTA 639
Db 421 GCTTCTCATGTGTGATTTTAACAGTGTGGATGATAAAGAGTGTCTCATATGTA 480
QY 640 AAGCATGATCTTGAATTAACCTGCTGCATTTTCTTCAACAGCCAGTTTCTGATGAG 699
Db 481 AAGCATGATCTTGAATTAACCTGCTGCATTTTCTTCAACAGCCAGTTTCTGATGAG 540
QY 700 AACAGGCTTTCAGTTTTCAGCTGGCATCATGTGTGTCAGATTTGCCAAATCAATTT 759
Db 541 AACAGGCTTTCAGTTTTCAGCTGGCATCATGTGTGTCAGATTTGCCAAATCAATTT 600
QY 760 GATTTGTTCTTTTACCATATCTT----- 784
Db 601 GATTTGTTCTTTTACCATATCTTAGTTTGAATTAATTAATAAAGTACACTGATG 660
QY 785 ----- 784
Db 661 GGCACTGTCTCTGTTTGGCTTGTGCTTTTCCCATGATGAGGACAGTCTAGTCTCAG 720
QY 785 ----- 784
Db 721 GGTAGTGAATAGTCTGTGATAGTATAGTACTAATAGTGAATATTAATTCTTACACAT 780
QY 785 ----- 784
Db 781 TGACTCAGACACACAGGTATGCACAACTTGTCTTTTGACCTAATACCTTTTACTTG 840
QY 785 ----- 784
Db 841 CTACTGTTCAATGACAAAACAGTGAACATCTGGCAATTTGACCTGGAACACTTTGCC 900
QY 785 -AGCAAGCGCAGAGAACATCAGCTGAGCAATTTTCCGAAGATTGGTCTAGAGAGTGC 843
Db 901 AAGCAAGCGCAGAGAACATCAGCTGAGCAATTTTCCGAAGATTGGTCTAGAGAGATG 960
QY 844 TCTCAACATGCTTTGTGCAAGATTTAAAGATCTTGTGATTTTCAAGATGATA 903
Db 961 TCTCAACATGCTTTGTGCAAGATTTAAAGATCTTGTGATTTTCAAGATGATA 1020
QY 904 ACATGATGGAAGAAAGCTGTGAATCTTCAAAAGAAAGTCTGGCTGATGATTTGAAA 963
Db 1021 ACATGATGGAAGAAAGCTGTGAATCTTCAAAAGAAAGTCTGGCTGATGATTTGAAA 1080
QY 964 TTGAATCTTGAAGACTGCTGATGAATTTTGAAGAAAGTCTGAGCAAGCAAGG 1023
Db 1081 TTGAATCTTGAAGACTGCTGATGAATTTTGAAGAAAGTCTGAGCAAGCAAGG 1140
QY 1141 TTAATATCTTCTTCAAGAAATCTGATGAATTTTATATGTCCAATPACTAGAGAACTTA 1200
Db 1084 TGAAGATCCGCTCATGCTGATGATGCTTATTCATATGAAAAGAAAGCAATGGAATTT 1143
Db 1201 TGAAGATCCGCTCATGCTGATGATGCTTATTCATATGAAAAGAAAGCAATGGAATTT 1260
QY 1144 GGATCAGCAAAAAGAAAC-GTACAAGTCCATGACAAATTTGTTCTTCCCTTC-AGGGT 1201
Db 1261 GGATCAGCAAAAAGAAACGTATCAAGTCCATGACAAATTTTCTTCCCTTCAGCGGG 1320
QY 1202 ACTTACACCAATATGAGCTGTGAAAATGGCCATCAATAGATGCTGAGACACACCAAAA 1261
Db 1321 CTTTACACCAATATGAGCTGTGAAAATGGCCATCAATAGATGCTGAGACACACCAAAA 1380
QY 1262 GTAAA 1266

PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249224P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-476159/51.
XX P-PSDB; ADM19864.
XX
PT Isolated nucleic acid molecule encoding a channel/transporter protein is
PS used in preventing, treating or ameliorating a medical condition.
XX
XX Claim 1; SEQ ID NO 192; 809pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule encoding a
CC channel/transporter protein or sequences at least 95% identical to a
CC these. The nucleic acids and proteins encoded by them are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. The antibodies to the proteins can also be used
CC in alleviating symptoms associated with the disorders and in diagnostic
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent
CC (ELISA). Disorders which are diagnosed or treated include autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC arrest, cerebrovascular disorders e.g. cerebral ischemia, angiodenesis,
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection.
CC The polypeptides can also be used to aid wound healing and epithelial
CC cell proliferation, to prevent skin aging due to sunburn, to maintain
CC organs before transplantation, for supporting cell culture of primary
CC tissues, to regenerate tissues and in chemotaxis. The polypeptides can
CC also be used as a food additive or preservative to increase or decrease
CC storage capabilities. This sequence corresponds to a gene of the
CC invention.
XX
SQ Sequence 1686 BP; 485 A; 348 C; 353 G; 500 T; 0 U; 0 Other;
Query Match 61.5%; Score 782; DB 5; Length 1686;
Best Local Similarity 77.3%; Fred. No. 3.3e-228;
Matches 1075; Conservative 0; Mismatches 180; Indels 135; Gaps 5;
QY 12 TCACCTGGCGGACGTGACCGGACCGCGCTTGAGGAGCGATCCCGCGC 71
DB 51 TCACCTGGCGGACGTGACCGGACCGCGCTTGAGGAGCGATCCCGCGC 110
QY 72 CCCCCTCTCTGACGAGCTGTTTCTTCAATAAAGACATGTAAGAACTGATTCACA 131
DB 111 CCCCCTCTCTGACGAGCTGTTTCTTCAATAAAGACATGTAAGAACTGATTCACA 170
QY 132 TTAGTATCATGTGACGATGTCATGCTGTGCTTCTTCTTCTTCTTCTGCTACT 191

DB 171 TTAGTATCATGTGACGATGTCATGCTGTGCTTCTTCTTCTTCTGCTACT 230
QY 192 TGTCTCTTGGACAAAACAATTCGCCCTGTACTGTTACGTGACTTTACTGAACGCCACAT 251
DB 231 TGTCTCTTGGACAAAACAATTCGCCCTGTACTGTTACGTGACTTTACTGAACGCCACAT 290
QY 252 TCTCCATGTAAGTTTATCATCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 311
DB 291 TCTCCATGTAAGTTTATCATCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 350
QY 312 ATTTTGGCATGCTGTTCAACAGATGATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 371
DB 351 ATTTTGGCATGCTGTTCAACAGATGATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 410
QY 372 ATGCTGGCAGTATGTAAGACAGCTAGTGAGCAGCTGCTGAGGCTTGGCAGTTTCCCA 431
DB 411 ATGCTGGCAGTATGTAAGACAGCTAGTGAGCAGCTGCTGAGGCTTGGCAGTTTCCCA 470
QY 432 GACTCCACGCTGTTGGCATGAGGGGAGCTGATGGAACCTGCTGTTTGGGAATGACAG 491
DB 471 GACTCCACGCTGTTGGCATGAGGGGAGCTGATGGAACCTGCTGTTTGGGAATGACAG 530
QY 492 TCATCAAAATATATAGATGCTGATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 551
DB 531 TCATCAAAATATATAGATGCTGATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 590
QY 552 TCTCTTAATGAAAGCTTCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 611
DB 591 TCTCTTAATGAAAGCTTCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 650
QY 612 GATTAATGAGTGTCTGCTGATGTAAGAAAGACATGATGATGATGATGATGATGATGATGAT 671
DB 651 GATTAATGAGTGTCTGCTGATGTAAGAAAGACATGATGATGATGATGATGATGATGATGAT 710
QY 672 TTTTCTTCAACAGCAGCTTCTGATGAGAAACAAGTCTTCAAGTTTTCAGTGGCATCA 731
DB 711 TTTTCTTCAACAGCAGCTTCTGATGAGAAACAAGTCTTCAAGTTTTCAGTGGCATCA 770
QY 732 TGTGTGAGGATGTCAGAGTCAAAATTTGATGTTTCTTAAACCATATCTTGAAG 791
DB 771 TGTGTGAGGATGTCAGAGTCAAAATTTGATGTTTCTTAAACCATATCTTGAAG 830
QY 792 CGCAGAAATCATGAGCTGAAGC-----AATT 818
DB 831 GAATTAATAATTAATAAGTACATGAGTGGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 890
QY 819 ACCGAAGATTGTCAGAGGAGTC----- 842
DB 891 TCCCATATGAGGACAGATGCTAGTCTCAGGCTCAGTGAATAGTCTGATATATATAT 950
QY 843 -----GTCACAAATG 854
DB 951 ACTAATATGAGAT 1010
QY 855 CTTTGTGCAAGA-----TTTAAAGATCTTGTGTATTTTCAAGATGAATTAAT 908
DB 1011 GCTTTGACCTATATACCTTTTAACTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1070
QY 909 GATGAAGAAGACGTTGATCTTAAAGAAAGTCTGCTGATGATTTGAAATTTGAA 968
DB 1071 TGGCAATTTTATCTTGAAGAAACCTTTTGCAGAGGACAGACAGAACTGAGTGAACAA 1130
QY 969 TCT--CTAGACCTGCTGATGTAAGTCTGAGAAATTTGAAGCTCAGAGCAAGGTTA 1026
DB 1131 TTTTACGAAGTTTGTGTCAGAGAGATGCTCAACATGCTGTTGTGCAAGATTTTAA 1190
QY 1027 AATCTCTTCTTCAAGAAATCTGATATATATATATATATATATATATATATATATATATAT 1086
DB 1191 GATCTTGTGTATTTTCAAGATGATATATATATATATATATATATATATATATATATATAT 1250
QY 1087 AA-----GATCGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1136

```

Db      1251 AAAGAAAGCTGGCTGATGATTTGAAAATTGATGGCTATTCAATATGAAAAGAGCAATG 1310
QY      1137 GAAATTTGATCAGCAAAAAGAAACGTACAAAGTCCCATGACAAATCTTGTCTTCTTCA 1196
Db      1311 GAAATTTGATCAGCAAAAAGAAACGTACAAAGTCCCATGACAAATCTTGTCTTCTTCA 1370
QY      1197 GCGGTACTTACACCAAAATAGGACTCTGAAAATGGCCATCAATAGATGGCTGAGACACAC 1256
Db      1371 GCGGTACTTACACCAAAATAGGACTCTGAAAATGGCCATCAATAGATGGCTGAGACACAC 1430
QY      1257 CAAAAGTAAA 1266
Db      1431 CAAAAGTAAA 1440

```

Search completed: February 5, 2005, 05:27:38
 Job time : 789 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: February 5, 2005, 08:48:01 ; Search time 696 Seconds
(without alignments)
3266.064 Million cell updates/sec

Title: US-10-077-111-13

Perfect score: 2047
Sequence: 1 MVXLIHTLADHGDDVNCACF.....LTPNRLKMAINRMLTTHQK 384

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 4390206 segs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL-frame.p2n.model -DEV-xlh
-Q=/cg2_1/USFPO.spool/US10077111/runat_04022005_121052_6794/app.query.fasta.1.583
-DB=N.GeneSeq_16Dec04 -QFM=fastcap -SUFFIX=p2n.fng -MINMATCH=0.1 -LOOPL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosome2 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=plc -NOM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USR=US10077111.QCN 1.1.644 @runat_04022005_121052_6794 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_GeneSeq_16Dec04.*

1: geneeqn1980s.*
2: geneeqn1990s.*
3: geneeqn2000s.*
4: geneeqn2001as.*
5: geneeqn2001bs.*
6: geneeqn2002as.*
7: geneeqn2002bs.*
8: geneeqn2003as.*
9: geneeqn2003bs.*
10: geneeqn2003cs.*
11: geneeqn2003ds.*
12: geneeqn2004as.*
13: geneeqn2004bs.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2047	100.0	1272	6	AAD45075 Human RET
2	1984	96.9	1553	6	AAD45071 Human RET
3	1984	96.9	1818	6	AAD45070 Human RET
4	1984	96.9	1996	5	ABV29028 Human pro
5	1984	96.9	1996	5	ABV23190 Human pro

6	1978	96.6	1773	8	ABZ24707 Human cel
7	1978	96.6	1811	10	ADB63535 Human cDN
8	1978	96.6	1817	10	ADB47507 Human cDN
9	1978	96.6	1844	4	AA158876 Human pol
10	1978	96.6	1844	5	ADG99098 DNA encod
11	1978	96.6	1844	9	ADB48858 Novel hum
12	1958	95.7	1908	6	AAD45076 Human RET
13	1854	90.6	1623	4	AA164914 Beta-Ltran
14	1853.5	90.5	1826	4	AA160662 Human pol
15	1715.5	83.8	1686	5	ADM19385 Novel hum
16	1616.5	79.9	1901	6	AAD45073 Mouse RET
17	1226.5	59.9	1281	4	AAFS8364 Human GTP
18	1000	48.9	838	5	ADM19628 Novel hum
19	717.5	35.1	668	13	ADD54461 Novel can
20	709	34.6	630	6	AAD45072 Human RET
21	627.5	30.7	446	5	ABV05038 Human pro
22	626.5	30.6	409	5	ABV14207 Human pro
23	621.5	30.4	441	5	ABV35301 Human pro
24	621.5	30.4	441	5	ABV44135 Human pro
25	597	29.2	366	4	AA136816 Probe #55
26	597	29.2	366	4	AAK30903 Human bon
27	597	29.2	366	4	AAK05311 Human bra
28	597	29.2	366	4	ABK30581 Human liv
29	597	29.2	366	6	ABK05651 Human gen
30	532	26.0	297	4	AA150015 Probe #18
31	532	26.0	297	4	AAK44005 Human bon
32	532	26.0	297	4	AAK18116 Human bra
33	532	26.0	297	4	ABK34659 Human liv
34	532	26.0	297	6	ABK18238 Human gen
35	509	24.9	419	3	AAH30188 Human col
36	352	17.2	520	6	AAD45074 Rat RET16
37	318	15.5	399	5	AAFE6462 Novel hum
38	308.5	15.1	466	4	ABA58107 Human foe
39	308.5	15.1	466	4	AA137710 Probe #63
40	308.5	15.1	466	4	AAK31840 Human bon
41	308.5	15.1	466	4	AAK06182 Human bra
42	308.5	15.1	466	4	ABK31528 Human liv
43	308.5	15.1	466	6	ABK06599 Human gen
44	300	14.7	180	4	ABA70710 Human foe
45	300	14.7	180	4	AA150879 Probe #19

ALIGNMENTS

RESULT 1	
AAD45075	
ID	AAD45075 standard; DNA; 1272 BP.
XX	
AC	AAD45075;
XX	
DT	27-DEC-2002 (first entry)
XX	
DE	Human RET16.2 splice variant DNA.
XX	
KW	Human; RET16; intracellular signal; inflammation-related disease; asthma;
KW	rheumatoid arthritis; psoriasis; multiple sclerosis; neuroprotective;
KW	transplant rejection; chronic obstructive pulmonary disease; TNF-alpha;
KW	inflammatory bowel disease; Crohn's disease; immunosuppressive; cancer;
KW	inacute respiratory distress syndrome; cardiact; ulcerative colitis;
KW	autoimmune disease; cystic fibrosis; gene therapy; cytostatic; neoplasm;
KW	cellular migration disorder; cell proliferation disorder; calcification;
KW	hypernatraemia; diabetes type 2; systemic lupus erythematosus; tumour;
KW	cardiovascular disease; Wegener's granulomatosis; atherosclerosis;
KW	thalassaemia; vasotropic; gene; ds.
XX	
OS	Homo sapiens.
XX	
FH	
FT	Key
CD	Location/Qualifiers
XX	111..1265
XX	/tag= a
XX	/product= "Human RET16.2 splice variant protein"
XX	
XX	W0200266494-A2.

XX		PD	29-AUG-2002.
XX		PF	15-FEB-2002; 2002MO-US005162.
XX		PR	16-FEB-2001; 2001US-0269366P.
XX		PT	29-MAY-2001; 2001US-0294181P.
XX		PA	(BRIM) BRISTOL-MYERS SQUIBB CO.
XX		PI	Todderud CG, Finger JN, Rillema J;
XX		DR	WPI; 2002-682760/73.
XX		P	P-PsDB; AAE28167.
XX		CC	New human, mouse or rat RET16 genes and proteins, involved in intracellular signaling cascade, useful for in gene therapy, particularly for treating e.g., ischemia, cystic fibrosis, autoimmune disease, cancers, tumors or neoplasms.
XX		SS	Claim 1; Page 166; 175pp; English.
XX		CS	The invention relates to human, mouse or rat RET16 genes and proteins, involved in intracellular signaling cascade. The RET16 protein or polynucleotide is useful for treating an inflammation-related disease or disorder, e.g., rheumatoid arthritis, juvenile arthritis, psoriasis, asthma, leukaemia-reperfusion, multiple sclerosis, rejection of organ or tissue transplants, chronic obstructive pulmonary disease, inflammatory bowel disease, Crohn's disease, ulcerative colitis, inacute respiratory distress syndrome, systemic lupus erythematosus, autoimmune disease, cystic fibrosis, cancer, tumours or neoplasms. This inflammation-related disease or disorder also includes disorders associated with aberrant activation of the TNF-alpha pathway, disorders associated with aberrant cellular migration, proliferation, metastasis, juvenile idiopathic arthritis, haematogenous metastases of tumour cells, hyperinflammation, diabetes type 2, atherosclerosis, cardiovascular disease, colon cancer, tumour progression, Wegener's granulomatosis, stem cell transplantation complications, leukaemia-reperfusion injury, thalassaemia, acute lung injury, graft rejection, ischaemic heart, coronary artery calcification or allergic inflammation. RET16 DNA is used in gene therapy. The present sequence is human RET16.2 splice variant DNA
SQ		Sequence	1272 BP; 349 A; 269 C; 290 G; 364 T; 0 U; 0 Other:
		Alignment Scores:	
		Pred. No.:	1,33e-191 Length: 1272
		Score:	2047.00 Matches: 384
		Percent Similarity:	100.00% Conservative: 0
		Best Local Similarity:	100.00% Mismatch: 0
		Query Match:	100.00% Indels: 0
		DB:	Gaps: 0
		US-10-077-111-J3 (1-384) x AAD45075 (1-1272)	
OY	1	MecVAllyvleuIllehIshThrLeuaAAsPHlSGLYASpApRYalaNcCYsVALaphe	20
Dd	111	ATGGGAAGAACGATTTCACACCATTAAGCTGCATGTGAGCAGATGTCAACTGCTGCTTIC	170
OY	21	SerPhesSerlleuMaIAshThrCyssSerleuAapRyShThIIeaRgeUryTsreLauAg	40
Dd	171	TCTTTTCCCCTTGCGCTACTGTGCTCTTGGAGCAAACAATTCGCCGTGAACTCGTTACGT	230
OY	41	AAPhetHrgIUleuProHisSerProLeuLySPheHisThrTYALaValHisCYsCVS	60
Dd	231	GACCTTAGCTGAAGCTGCCAACATTTCTCATTTGAAGTTTCAATACCATACTGTCACCTGTC	290
OY	61	CysPheSerProSeRGylYHIsileLeualaaserCySserThraSpglYthrThrVaalleu	80
Dd	291	TGTTTTCCCCCTTCAGAGCATATTTTGGCATCGTGTCCAAGATGTAACAAGTGTCACTGTCTTA	350
OY	81	TrpaSenHrgIUanAngIGylImetLeualavalmetGluglnProseRgylSerProval	100
Dd	351	TGGAATACTGAAAAGGACAGATGTGGGAGAGATGGAAACAGCTTAGTGGAAGCCCTGTG	410

QY	101	ArgValCYGslPbSerProApSerThrCysLeuAlaSerGlyAlaAlaAPGlyThr	120
Db	411	AGGGTTTGGCAAGTTTCCCAAGATCCACGCTTTGGCATCGGGGAGCTGATGAACT	470
QY	121	ValValLeuTrpAsnAlaGlnSerTyrLeuLeuTyrArgCysGlySerValLysAspGly	140
Db	471	GTGGTTTGTGGAAATGCACAGCTACATACAAATTATATAGATGTGGTAGTGTAAAGATGGC	530
QY	141	SerLeuAlaAlaCysAlaPbSerProAsnGlySerPhePheValThrGlySerSerCys	160
Db	531	TCTTTGGCGGCAATGGTCATTTTCTCTAATGAAAGCTTCTTGTGACCTGGCTCTCATGT	590
QY	161	GlyAspLeuThrValTrpAspArgPheSerArgCysLeuHisSerGlyLysAlaHisAsp	180
Db	591	GGGATTTAAACAAGTGTGGATGATTAATATGAGGCTGTGCTAGTATGAAAAAGCATGAT	650
QY	181	LeuGlyIleThrCysCysAspPbSerSerGlnProValSerAspGlyGlnGlnGlyLeu	200
Db	651	CTTGAAATTAACCTGCTGGATTTTCTTCTTCAACAGCCAGTTTCTGATGAGAAACAAGGCTT	710
QY	201	GlnPhePheArgLeuAlaSerCysGlyGlnAspCysGlnValLysIleTrpIleValSer	220
Db	711	CAATTTTGTTCGACTGGCAATGAGTGTCAAGATTTGGCCAAAGTCAAAATTTGGATGTTTTCT	770
QY	221	PheThrHisIleLeuAlaArgArgThrGlnHisGlnLeuLysGlnPheThrGluAspTrp	240
Db	771	TTTATCCCATATCTTTGCAGAGGCGCACAGAACATACAGCTGAAGCATTTTACCGAAAGTTGG	830
QY	241	SerGlnGlnValValSerThrTrpLeuCysAlaGlnAspLeuLysAspLeuValGlyIle	260
Db	831	TCAGAGAGAGTGTGTCAACATGGCTTTGTGCACAAATTTAAAGATCTTGTGGTATT	890
QY	261	PheLysMetAsnAsnIleAspGlyLysGlnLeuLeuLeuLeuThrLysGlnSerLeuAla	280
Db	891	TTCAATATTAATAACATTTGATGAGAAAGAACTGTTGATATCTTCAAAAGAAAGTCTGGCT	950
QY	281	AspAspLeuLysIleGlnSerLeuGlyLeuArgSerLysValLeuArgLysIleGlnGln	300
Db	951	GATGATTTGAAATTTGAAATCTCTAGAGCTGGTGTGTAATGCTGTAGAGAAATTTGAAGAG	1011
QY	301	LeuArgThrLysValLysSerLeuSerSerGlyIleProAspGlnPheIleCysProIle	320
Db	1011	CTCAGAGCCAAAGTTTAAATCCCTTTCTTCAGAAATTCCTATGAATTTATATGTCCAAATA	1071
QY	321	ThrArgGlnLeuMetLysAspProValIleLeuSerAspGlyLysSerTyrGlnLysGln	340
Db	1071	ACTAGGAATCTTATATAAAGATCCGGTCACTGCATCAAGATGCTATTTCAATATGAAAAAGAA	1131
QY	341	AlaMetGluAsnTrpIleSerLysLysLysArgThrSerProMetThrAsnLeuValLeu	360
Db	1131	GCATGGAAGAAATTTGATACGCAAAAAAGAAAGCTACAGTCCCATGACAAATCTGTGTTCTT	1191
QY	361	ProSerAlaValLeuThrProAsnArgThrLeuLysMetAlaIleAsnArgTrpLeuGln	380
Db	1191	CCTTCAAGCGGTACTTACACCAAAATAGGACTCTGAAATATGGCCATCATATGATGGCTGAGAG	1251
QY	381	ThrHisGlnLys	384
Db	1251	ACACACCAAAAG	1262
RESULT 2			
AADA5071			
ID	AADA5071	standard; DNA; 1553 BP.	
AADA5071;			
XX	27-DEC-2002	(first entry)	
XX	Human REF16	open reading frame (ORF) DNA.	
XX	Human; REF16;	intracellular signal; inflammation-related disease; asthma;	
XX	rheumatoid arthritis; psoriasis;	multiple sclerosis; neuroprotective;	

XX	transplant rejection; chronic obstructive pulmonary disease; TNF-alpha;
KW	inflammatory bowel disease; Crohn's disease; immunosuppressive; cancer;
KM	inacue respiratory distress syndrome; cardiact; ulcerative colitis;
KM	autoimmune disease; cystic fibrosis; gene therapy; cyostatic; neoplaem;
KW	cellular migration disorder; cell proliferation disorder; calcification;
KW	hyperinsulinemia; diabetes type 2; systemic lupus erythematosus; tumour;
KW	cardiovascular disease; Wegener's granulomatosis; atherosclerosis;
KW	thalassaemia; vasotropic; open reading frame; ORF; gene; ds.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	CDS
FT	3. 1541
FT	/*cag= a
FT	/product= "Human RET16 protein"
FT	/note= "CDS does not include start codon"
FT	/partial
XX	
PN	MO200266494-A2.
XX	
PD	29-AUG-2002.
XX	
XX	15-FEB-2002; 2002MO-US005162.
PP	16-FEB-2001; 2001US-0269366P.
XX	
PR	29-MAY-2001; 2001US-0294181P.
XX	
PA	(BRIM) BRISTOL-MYERS SQUIBB CO.
XX	
PI	Toddernud CG, Finger JN, Rillema J;
XX	
DR	WPI: 2002-682760/73.
XX	
PT	P-PSDB: AAE28164.
PT	
XX	New human, mouse or rat RET16 genes and proteins, involved in
XX	intracellular signaling cascade, useful for in gene therapy, particularly
XX	for treating e.g. ischemia, cystic fibrosis, autoimmune disease, cancers,
XX	tumors or neoplasms.
XX	
PS	Claim 1; Page 150; 175bp; English.
XX	
CC	The invention relates to human, mouse or rat RET16 genes and proteins,
CC	involved in intracellular signaling cascade. The RET16 protein or
CC	polynucleotide is useful for treating an inflammation-related disease or
CC	disorder, e.g. rheumatoid arthritis, juvenile arthritis, psoriasis,
CC	asthma, ischemia-reperfusion, multiple sclerosis, rejection of organ or
CC	tissue transplants, chronic obstructive pulmonary disease, inflammatory
CC	bowel disease, Crohn's disease, ulcerative colitis, inacute respiratory
CC	distress syndrome, systemic lupus erythematosus, autoimmune disease,
CC	cystic fibrosis, cancers, tumours or neoplasms. This inflammation-related
CC	disease or disorder also includes disorders associated with aberrant
CC	activation of the TNF-alpha pathway, disorders associated with aberrant
CC	cellular migration, proliferation, metastasis, juvenile idiopathic
CC	arthritis, haemoglobin metastases of tumour cells, hyperinsulinemia,
CC	diabetes type 2, atherosclerosis, cardiovascular disease, colon cancer,
CC	tumour progression, Wegener's granulomatosis, stem cell transplantation
CC	complications, ischemia-reperfusion injury, thalassaemia, acute lung
CC	injury, graft rejection, ischaemic heart, coronary artery calcification
CC	or allergic inflammation. RET16 DNA is used in gene therapy. The present
CC	sequence is human RET16 open reading frame (ORF) DNA
XX	
XX	
SO	Sequence 1553 BP; 428 A; 327 C; 348 G; 450 T; 0 U; 0 Other;
XX	
XX	Alignment Scores:
XX	Pred. No.: 2.86e-185 Length: 1553
XX	Score: 1984.00 Matches: 383
XX	Percent Similarity: 80.46% Conservative: 0
XX	Best Local Similarity: 80.46% Mismatches: 1
XX	Query Match: 96.92% Indels: 92
XX	DB: 6 Gaps: 1

QY	1	MetValValLeuIleHisIleThrLeuAlaAspHisGlyAspAspValAlaSerCysValAlaPhe	20
Db	111	ATGGTAACACTGCATTACACATTACTATCATGTGACGATGTCAACTGCTGTGCTTC	170
QY	21	SerPheSerLeuLeuAlaThrCysSerLeuAspLysTrpIleArgLeuTyrSerLeuArg	40
Db	171	TCCTTTTCCCTGGCTGACTTGGTCCCTTGGCAAAACAATTGGCTGTGACTGGTTACGT	230
QY	41	AspPheThrGluLeuProHisSerProLeuLysPheIleTrpTyrAlaValHisCysCys	60
Db	231	GACTTACTGAACTGCCACATTCCTCCATTTGAAATTCATTACTTACCTGATGCTCCACTGCTGC	290
QY	61	CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrThrValLeu	80
Db	291	TGTTTCTCCCTTCAAGACATATTTTGGCATCTGTGTTACACAGATGTACCACTGTCTTA	350
QY	81	TrpAsnThrGluAsnGlyIleMetLeuAlaValMetGluInProSerGlySerProVal	100
Db	351	TGGATACTGAAATATGACAGATGTGGCAGATGAAACAGCTAGTGGCAGCCCTGTG	410
QY	101	ArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyAlaAlaAspGlyThr	120
Db	411	AGGGTTTGGCCAGTTTCCCCAGACTCCACGTGTTGGCATCAGGGGGCAGCTGATGAACCT	470
QY	121	ValValLeuTrpAsnAlaGlnSerClyLeuTyrLeuTyrArgCysGlySerValLysAspGly	140
Db	471	GTGGTTTGTGGAATGACACAGCTCATACAAATTAATATAGATGTGTGTATGAAGATGGC	530
QY	141	SerLeuAlaAlaCysAlaPheSerProAsnGlySerPhePheValThrGlySerSerCys	160
Db	531	TCCTTTGGGGGAGTGTGCATTTTCTCCTAATGGAAGCTTCTTTGTACCTGGCTCCCTCATGT	590
QY	161	GlyAspLeuThrValTrpAspAspLysMetArgCysLeuHisSerGluLysAlaHisAsp	180
Db	591	GGTGATTTAAACAGTGGGATGATTAATATGAGTGTCTGCATATGGAAGAAAAGCATGAT	650
QY	181	LeuGlyIleThrCysCysAspPheSerSerGlnProValSerAspGlyGluGlnLysLeu	200
Db	651	CTTGAATATACCTGCTGGCATTATTTCTTCACAGCCAGTTTCTGATGGAGAACACAGGCTCTT	710
QY	201	GlnPhePheArgLeuAlaSerCysGlyGlnAspCysGlnValLysIleTrpIleValSer	220
Db	711	CAGTTTTCGACCTGGCATCATGTGTGTCAGAGATTGCCAAGTCACAAATTTGGATTGTCT	770
QY	221	PheThrHisIleLeu-----	225
Db	771	TTTACCCTATCTTAGGTTTGAATTAATAATAAAGTACACTGAGTGGGCACTGTGCT	830
QY	225	-----	225
Db	831	CCTGTTCGCTGTGTGCTTTTCCCGTATGGGAGATGATCTCAAGGTCACTGAT	890
QY	225	-----	225
Db	891	AAATCTGTCATATATATATGATATCTAATATCTAGAAATATATCTTACACACTTGACTACAGC	950
QY	225	-----	225
Db	951	ACCAAGTATGTCACAACCTTGTGCTTTGGACCTTAATACCTTTTACTTGCTACTGTTCA	1011
QY	226	-----AlaArgArg	228
Db	1011	ATGGACAAACAGTGAACATCTGGCAATTTGACCTGGAACACTTTTGCCAAAGGGCGC	1070
QY	229	ThrGluHisGlnLeuLysGlnPheThrGluAspTrpSerGluGluValValSerThrTrp	248
Db	1071	ACGAAACATCACTGAGAACCAATTTACGAAATTTGTGTAGAGAGAGATGTCTCAACATGG	1133
QY	249	LeuCysAlaGlnAspLeuLysAspLeuValGlyIlePheLysMetAsnAsnIleAspGly	268
Db	1131	CTTTGTGTGACAAAGATTTAAAGATCTTGTGTATTTTCAAGATGAATATACCTTGATGGA	1199
QY	269	LysGluLeuLeuAsnLeuThrLysGluSerLeuAlaAspAspLeuLysIleGluSerLeu	288

Db 1191 AAGAACTGTGATCTTACAAAAGAGTGGCTGATGATTTGAAATTTGAAATCTCTA 1250
Qy 289 G1YleuAserSerValIleuAArgLysIleG1uG1uLeuAArgThryLysValIysSerLeu 308
Db 1251 GGACTGCTGTGTAAGAGCTGAGGAAATTTGAAAGAGTCCAGGACCAAGTTAAATCCCTT 1310
Qy 309 SerSerG1YIleProAArgIupheIleCyseProIleThraArgIuLeuMetLysAsePro 328
Db 1311 TCTTCAGGAATTCCTGATGAATTTATATGTCCAATACTAGAGAACTTATGAAAGATCCG 1370
Qy 329 ValIleAlaSerArgIYTYrSerTYrG1uLysG1uAlaMetG1uAantP1IleSerLys 348
Db 1371 GTCATGCGCATCAATGCGTATTCATATGAAAGAACCAATGGAAATTGATCAGCAAA 1430
Qy 349 LysIleAArgThrySerProMetThraAsnLeuValIleuProSerAlaValIleuThryProAsn 368
Db 1431 AAGAAAGCTGACAGTCCCATGACAAATCTGTTCTTCTTCAGCGGACTTACCAAAAT 1490
Qy 369 ArgThryLeuLysMetAlaIleAsnArgTYrLeuG1uThryIseG1uLys 384
Db 1491 AGGACTCTGAATAATGGCCATCAATGATGCTGAGACACACCAAG 1538

RESULT 3
AAD45070 standard; DNA; 1818 BP.
AAD45070;
27-DEC-2002 (first entry)

Human RET16 DNA.
Human RET16.
Human; RET16; intracellular signal; inflammation-related disease; asthma;
rheumatoid arthritis; psoriasis; multiple sclerosis; neuroprotective;
transplant rejection; chronic obstructive pulmonary disease; TNF-alpha;
inflammatory bowel disease; Crohn's disease; immunosuppressive; cancer;
inacute respiratory distress syndrome; cardiac; ulcerative colitis;
autoimmune disease; cystic fibrosis; gene therapy; cytotoxic; neoplasm;
cellular migration disorder; cell proliferation disorder; calcification;
hypernatremia; diabetes type 2; systemic lupus erythematosus; tumour;
cardiovascular disease; Wegener's granulomatosis; atherosclerosis;
thalassaemia; vasotropic; gene; db.

Homo sapiens.
Key Location/Qualifiers
FT CDS 148..1578
FT /tag= a
FT /product= "Human RET16 protein"

MO200266494-A2.
29-AUG-2002.
15-FEB-2002; 2002MO-US005162.
16-FEB-2001; 2001US-0269366P.
29-MAY-2001; 2001US-0294181P.
(BRIM) BRISTOL-MYERS SQUIBB CO.
Todderud CG, Finger JN, Rillema J;
WPI; 2002-682760/73.
P-PSDB; AAE28163.

New human, mouse or rat RET16 genes and proteins, involved in
intracellular signaling cascade, useful for in gene therapy, particularly
for treating e.g. ischemia, cystic fibrosis, autoimmune disease, cancers,
tumors or neoplasms.

Claim 1; Page 147-148; 175pp; English.

XX The invention relates to human, mouse or rat RET16 genes and proteins,
CC involved in intracellular signaling cascade. The RET16 protein or
CC polynucleotide is useful for treating an inflammation-related disease or
CC disorder, e.g. rheumatoid arthritis, juvenile arthritis, psoriasis,
CC asthma, ischaemia-reperfusion, multiple sclerosis, rejection of organ or
CC tissue transplants, chronic obstructive pulmonary disease, inflammatory
CC bowel disease, Crohn's disease, ulcerative colitis, inacute respiratory
CC distress syndrome, systemic lupus erythematosus, autoimmune disease,
CC cystic fibrosis, cancers, tumours or neoplasms. This inflammation-related
CC disease or disorder also includes disorders associated with aberrant
CC activation of the TNF-alpha pathway, disorders associated with aberrant
CC cellular migration, proliferation, metastasis, juvenile idiopathic
CC arthritis, haematogenous metastases of tumour cells, hypernatremia,
CC diabetes type 2, atherosclerosis, cardiovascular disease, colon cancer,
CC tumour progression, Wegener's granulomatosis, stem cell transplantation
CC complications, ischaemia-reperfusion injury, thalassaemia, acute lung
CC injury, graft rejection, ischaemic heart, coronary artery calcification
CC or allergic inflammation. RET16 DNA is used in gene therapy. The present
CC sequence is human RET16 DNA

Sequence 1818 BP; 523 A; 373 C; 382 G; 540 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 3,58e-185 Length: 1818
Score: 1984.00 Matches: 383
Percent Similarity: 80.46% Conservative: 0
Best Local Similarity: 80.46% Mismatches: 1
Query Match: 96.92% Indels: 92
DB: 6 Gaps: 1

US-10-077-111-13 (1-384) x AAD45070 (1-1818)

Qy 1 MetValIleLeuIleHisThryLeuAlaAspHisG1YAspAspValAsnCysCysAlaPhe 20
Db 148 ATGTGAAACTGATTTTCACATTAGCTGATCAGTGAGATGTCACATGCTGTGCTTC 207
Qy 21 SerPheSerLeuLeuAlaThryCysSerLeuAspLysThryIleArgLeuThrySerLeuArg 40
Db 208 TCTTTTCCCTCTTGCTACTTCTCTCTGACAAACAATTCGCTGACTGATTAACG 267
Qy 41 AspPheThryG1uLeuProHisSerProLeuLysPheHisThryTrpAlaValHisCysCys 60
Db 268 GACTTTACTGACATCGCACATTTCTCATTTGAATTTTCATACCTATCTGCTCCACTGCTG 327
Qy 61 CysPheSerProSerG1YHisIleLeuAlaSerCysSerThryAspG1YThryValIleu 80
Db 328 TGTTTCTCCCTTCAGGACATATTTTGGCATGCTGTTCAACAGATGCTACACTGCTCTA 387
Qy 81 TrpAsnThryG1uAnG1YIleMetLeuAlaValMetG1uG1uProSerG1YSerProVal 100
Db 388 TGGATATCTGAAATGAGACAGATGCTGGCAGTGATGGAACAGGCTAGTGAGCCCTGTG 447
Qy 101 ArgValCysG1uInPheSerProAspSerThryCysLeuAlaSerG1YValAlaAspG1YThr 120
Db 448 AGGGTTTGCAGATGTTTCCCGACCTCCAGCTGTTGGCATCAGAGGCGCAGCTGATGAACT 507
Qy 121 ValIleLeuThryAsnAlaG1uInSerTYrLysLeuTYrArgCysG1YSerValIysAspG1Y 140
Db 508 GTGGTTTGTGGAATGACAGTCATACAAATTTATAGATGTGTGTGTTAAAGATGGC 567
Qy 141 SerLeuAlaAlaCysAlaPheSerProAsnG1YSerPheValThryG1YSerSerCys 160
Db 568 TCTTGCGCGCATGTCATTTCTCTTAATGGAAGCTTCTTGTCTGCTGCTCCTCATGT 627
Qy 161 G1YAspLeuThryValIleAspAspLysMetArgCysLeuHisSerG1YValAlaHisAsp 180
Db 628 GGTGATTTTAACTGTGTGGATGTATTAATAAGAGTGTCTGATGTATTAATAAGACATGAT 687
Qy 181 LeuG1YIleThryCysAspPheSerSerG1uProValSerAspG1YG1uG1uLysLeu 200
Db 688 CTGGAATTTACCTGCTGCGATTTTCTTCAACAGCAGTTTCTGATGAGAACAGGCTTT 747

QY 201 GlnPhePheArgLeuAlaSerCysGlyGlnAspCysGlnValLysIleTyrIleValSer 220
Db 748 CAGTTTTCGACCTGGCATCATGTGTCAGATTGGCCAAAGTCAAAATTTGGATTGTTCT 807
QY 221 PheThrIleIleLeu----- 225
Db 808 TTAAACCAATCTTAGGTTTGAATTAAATATATAAGTACAGAGTGGGCACTGTGCT 867
QY 225 ----- 225
Db 868 CCTGTTCTGGCTTGTCTTTTCCCATGATGGGAGATGCTAGTCTCAGGCTCAGTGAT 927
QY 225 ----- 225
Db 928 AAGTCTGTCATAGTATATGATCTAATACGAGAATATCTTCAACATTTGACTCAGCAC 987
QY 225 ----- 225
Db 988 ACCAGTATGTCAACAATTGTCTTTGGCACTTAATACCTTTTACTTGTCTACTGTCTCA 1047
QY 226 -----AlaArgArg 228
Db 1048 ATGCACAAACAGTGAACATTCGTGCATTTGACCTGGAACACTTTGCCAAGCAAGGCGC 1107
QY 229 ThrGlnIleGlnLeuLysGlnPheThrGlnAspTyrSerGlnGlnValIleSerThrTyr 248
Db 1108 ACAGAACATCAGCTGGAAGCAATTTTACCAAGATTGCTCAGAGGAGCATGTCTCAACATCG 1167
QY 249 LeuCyAlaGlnAspLeuLysAspLeuValGlyIlePheLysMetAsnAsnIleAspGly 268
Db 1168 CTTGTGTCAACAAGTTTAAAGATCTTGTGTGTTTTCATGAATGAATTAACATTGATGGA 1227
QY 269 LysGlnLeuLeuAsnLeuThrLysGlnSerLeuAlaAspAspLeuLysIleGlnSerLeu 288
Db 1228 AAAAAGACTGTGAATCTTCAAAAGAAAGTGTGCTGATGATTTGAATAATTCCTCTA 1287
QY 289 GlyLeuArgSerLysValIleuArgLysIleGlnGlnLeuArgThrLysValLysSerLeu 308
Db 1288 GGACTGCGTACTTAAGTCTGAGAAATTAAGAGCTCAGAGCAACAGTTAAATCCCTT 1347
QY 309 SerSerGlyIleProAspGlnPheIleCysProIleThrArgLysLeuMetLysAspPro 328
Db 1348 TCTTCAGGAATTCCTGATGATTTATATGTCCTCAATTAAGTAAGAACTTATGAAAGATCCG 1407
QY 329 ValIleAlaSerAspGlyTyrSerTyrGlnLysGlnAlaMetGlnAsnTyrIleSerLys 348
Db 1408 GTCATCCCATCAGATGGCTATTCATATGAAAGAAAGCAATGGAATTTGATCAGCAMA 1467
QY 349 LysLysArgThrSerProMetThrAsnLeuValLeuProSerAlaValLeuThrProAsn 368
Db 1468 AAGAAAGCTACAGTCCCATCAACAAATCTTGTCTTCTTCAAGCGGATCTTACACCAAT 1527
QY 369 ArgThrLeuLysMetAlaIleAsnArgTyrLeuGlnThrHisGlnLys 384
Db 1528 AGGACTCTGAAATATGGCCATCATATGATGGCTGAGAGCACACCAAAAG 1575

RESULT 4
ABV29028
ID ABV29028 standard; cDNA; 1996 BP.
XX
XX ABV29028;
AC
XX 16-SEP-2002 (first entry)
DT
XX
XX Human prostate expression marker cDNA 29019.
DE
XX
XX Human: prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200160860-A2.
XX

PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001MO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA
PI Schlegel R, Endege WO, Monahan JB;
PI WPI; 2001-662795/76.
DR
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 6145-6146; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 1996 BP; 615 A; 378 C; 447 G; 543 T; 0 U; 13 Other;
SQ
Alignment Scores:
Pred. No: 4.09e-185 Length: 1996
Score: 1984.00 Matches: 383
Percent Similarity: 80.46% Conservative: 0
Best Local Similarity: 80.46% Mismatches: 1
Query Match: 96.92% Indels: 92
DB: Gaps: 1
US-10-077-111-13 (1-384) x ABV29028 (1-1996)

QY 1 MetValLysLeuIleHisThrLeuAlaAspHisGlyAspAspValAencCysAlaPhe 20
Db 16 ATGCTGAACCTGATTCACACATTAGCTGATCAAGTACAGATGTCATCTGTGCTTTC 75
QY 21 SerPheSerLeuLeuAlaThrCysSerLeuAspLysThrIleArgLeuTyrSerLeuArg 40
Db 76 TCCTTTTCCTCTTGGCTACTTGTCTGTCGACAAAACAATTTGGCTGATCGTTACGT 135
QY 41 AspPheThrGlnLeuProHisSerProLeuLysPheHisThrTyrAlaValHisCysGys 60
Db 136 GACTTTACTGAACTGACACATTCCTCCATTAAGTTTCATCTCATGTGCTGCCACTGTGC 195
QY 61 CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyTyrThrValLeu 80
Db 196 TGTTCCTCCCTTCAGGACATATTTTGGCATCTGTTCAACAATGTATCCACTGTCTTA 255
QY 81 TrpAsnThrGlnAsnGlnGlnMetLeuAlaValMetGlnGlnProSerGlySerProVal 100
Db 256 TGGATATCTGAAAATGACAGATGCTGGCAGTGTGGAACAGCTAGTGGCAGCCCTGTG 315
QY 101 ArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyValAlaAspGlyThr 120
Db 316 AGGCTTTGCCAGTTTCCCAAGCTCCACCTGTTTGGCATCAGGGGCGCATGTGAACT 375
QY 121 ValValLeuThrPheAsnAlaGlnSerTyrLysLeuTyrArgCysGlnSerValLysAspGly 140

```
Db      ||| 376 GTGGTTTGGATGACAGTCATACAAATTAATAGATGTGTAAAGATGGC 435
Qy      ||| 141 SerLeuAlaIaIaCysAlaPheSerProAsnGlySerPheValThrGlySerSerCys 160
Db      ||| 436 TCCTTGGCGGCAATGATTTTCTCCAAATGAAAGCTCTTTGCTGCTGCTCTCATGT 495
Qy      ||| 161 GlyAspLeuThrValITrapAspAspLysMetArgCysLeuHisSerGlyValAsp 180
Db      ||| 496 GGTGATTTAAACAGTGTGGATGATTAATAAGAGGTGTCTGCATAGTGAATAAGCAATGAT 555
Qy      ||| 181 LeuGlyIleThrCysCysAspPheSerSerGlnProValSerAspGlyValGlnGlyLeu 200
Db      ||| 556 CTGGAAATTACCTGCTGCGATTTTCTTCAACAGCAATTTCTAATGAGAAACAAGCTTT 615
Qy      ||| 201 GlnPhePheArgLeuAlaSerCysGlyValAspCysGlnValIleValIleValSer 220
Db      ||| 616 CAGTTTTCGACTGGCATATGTGGTCAAGATTGCCAAGTCAAAATTTGGATTGTTCT 675
Qy      ||| 221 PheThrHisIleLeu----- 225
Db      ||| 676 TTACCATATCTTAGGTTTGAATTAAATATATAAGTACACTGAGTGGGCACTGTGCT 725
Qy      ||| 225 ----- 225
Db      ||| 736 CCGTTCTGGCTGTGCTTTTCCCATGATGGGAGATGCTAGTCAAGGTCAGTGAT 795
Qy      ||| 225 ----- 225
Db      ||| 796 AAGTCTGTCTATATATGATTAATCTAATCTGAGAAATTAATTCTACACATTGACTCAGCAC 855
Qy      ||| 225 ----- 225
Db      ||| 856 ACCAGGATGTGCACAACTTGCTTTTGCACTTAATACCTTTTACTTGCTAGTGTCA 915
Qy      ||| 226 -----AlaArgArg 228
Db      ||| 916 ATGACAAAACAGTGAAATCTGGCAATTTGACCTGGAAACACTTTCACAAAGAGCGC 975
Qy      ||| 229 ThrGluHisGlnLeuLysGlnPheThrGluAspTrpSerGluGluValValSerThrTrp 248
Db      ||| 976 ACGAATCATCAGCTGAAAGCAATTTACCGAAGATTGGCAAGAGGAGCATGTCTCAACATGG 1035
Qy      ||| 249 LeuCysAlaGlnAspLeuLysAspLeuValGlyIlePheLysMetAsnAsnIleAspGly 268
Db      ||| 1036 CTTTGTGCACAAAGTTTAAAGAACTTGTGTGATTTTCAAGATGATTAATGATGAGA 1095
Qy      ||| 269 LysGluLeuLeuAsnLeuThrLysGluSerLeuAlaAspAspLeuLysIleGluSerLeu 288
Db      ||| 1096 AAAGAACTGTGAATCTTACAAAGAAAGTGGCTGATGATTTGAAATTTGAATCTCTA 1155
Qy      ||| 289 GlyLeuArgSerLysValLeuArgLysIleGluGluLeuArgThrLysValLysSerLeu 308
Db      ||| 1156 GGAAGTGGTGAAGAGTGAAGGAAATTAAGAAAGTCAAGAACAAAGTTAAATCCCTT 1215
Qy      ||| 309 SerSerGlyIleProAspGluPheIleCysProIleThrArgGluLeuMetLysAspPro 328
Db      ||| 1216 TCTTCAGGAATTCCTGATGATTTATATGTCCTCAATACAGAACTTAAGAAAGATCCG 1275
Qy      ||| 329 ValIleAlaSerAspGlyIleLysSerTyrGluLysGluAlaMetGluAsnTrpIleSerLys 348
Db      ||| 1276 GTCATCGCATCAGATGGCTATTTCATATGAAAGAAAGCAATGGAATTTGGATCAGCAA 1335
Qy      ||| 349 LysLysArgThrSerProMetThrAsnLeuValLeuProSerAlaValIleLeuThrProAsn 368
Db      ||| 1336 AAAGAAAGTACAAATCCCATGACAAATCTGTCTTCTTCAAGGAGTACTTAACCAAAAT 1395
Qy      ||| 369 ArgThrLeuLysMetAlaIleAsnArgTrpLeuGlnThrHisGlnLys 384
Db      ||| 1396 AGAAGCTGAAGATGSCCATCAATAGATGGCTGAGACACACCAAAAG 1443
```

```
Id      ABV23190 standard; cDNA; 1996 BP.
Xx      AC      ABV23190;
Xx      DT      16-SEP-2002 (first entry)
Xx      DE      Human prostate expression marker cDNA 23181.
Xx      KW      Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
Xx      KW      pharmacogenomic marker; gene; ss.
Xx      OS      Homo sapiens.
Xx      PN      WO200160860-A2.
Xx      PD      23-AUG-2001.
Xx      PF      20-FEB-2001; 2001WO-US005171.
Xx      PR      17-FEB-2000; 2000US-0183119P.
Xx      PR      16-MAR-2000; 2000US-0189862P.
Xx      PR      25-MAY-2000; 2000US-0207454P.
Xx      PR      09-JUN-2000; 2000US-0211314P.
Xx      PR      18-JUL-2000; 2000US-0219807P.
Xx      PR      13-DEC-2000; 2000US-0235281P.
Xx      PA      (MIL-) MILLENNIUM PREDICTIVE MEDICINE INC.
Xx      PI      Schlegel R, Endege WO, Monahan JB;
Xx      DR      WPI, 2001-662795/76.
Xx      PT      Novel isolated nucleic acid molecule associated with cancerous state of
Xx      PT      prostate cells and correlating with presence of prostate cancer, useful
Xx      PT      for detecting presence of prostate cancer, stage of prostate cancer.
Xx      PS      Claim 1; Page 4166-4167; 11750pp; English.
Xx      CC      The invention relates to an isolated nucleic acid molecule (1) comprising
Xx      CC      a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
Xx      CC      specification or its complement. (1) is useful for: (a) assessing whether
Xx      CC      a patient is afflicted with prostate cancer; (b) monitoring the
Xx      CC      progression of prostate cancer in a patient; (c) assessing the efficacy
Xx      CC      of a test compound to inhibit prostate cancer in a patient; (d) assessing
Xx      CC      the efficacy of a therapy for inhibiting prostate cancer in a patient;
Xx      CC      (e) selecting a composition for inhibiting prostate cancer in a patient;
Xx      CC      (f) assessing the prostate cell carcinogenic potential of a compound; (g)
Xx      CC      determining whether prostate cancer has metastasized in a patient; (h)
Xx      CC      assessing the aggressiveness or indolence of prostate cancer in a patient
Xx      CC      ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
Xx      SQ      Sequence 1996 BP; 615 A; 378 C; 447 G; 543 T; 0 U; 13 Other;

Alignment Scores:
Pred. No.: 4,09e-185 Length: 1996
Score: 1984.00 Matches: 383
Percent Similarity: 80.46% Conservative: 0
Best Local Similarity: 80.46% Mismatches: 1
Query Match: 96.92% Indels: 92
DB: 5 Gaps: 1

US-10-077-111-13 (1-384) x ABV23190 (1-1996)

Qy      1 MetValLysLeuLeuIleHisThrLeuAlaAspHisGlyAspAspValAsnCysCysAlaPhe 20
Db      16 ATGGTAAACTGATTACACATTAGCTGATCAATGAGAGATGCAACGCTGTGCTTC 75
Qy      21 SerPheSerLeuLeuAlaThrCysSerLeuAspLysThrIleArgLeuTyrSerLeuArg 40
Db      76 TCCTTTTCCTCTTGGCTACTTCTCTCTTGACAAACAATTCCTGACTGTTACGT 135
Qy      41 AspPheThrGluLeuProHisSerProLeuLysPheHisThrTyrAlaValHisCysCys 60
```


Db 136 GACTTACTGAACTGCCACATTCCTCATTTGAAAGTTTCATACCTATGCTGCTCCACTGCTGC 195
 Qy 61 CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrValLeu 80
 Db 196 TGTTCCTCCCTTCCAGACATATTTGGCATCGTGTCACAGATGGACACAGCTCTTA 255
 Qy 81 TrpAsnThrGluAsnGlyGlnMetLeuAlaValMetGluGlnProSerGlySerProVal 100
 Db 256 TGGAAATCTGAAAATGAGCAGATCTGCGACAGTGAAGAACGCTAGTGGAGCCCTGTG 315
 Qy 101 ArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyAlaAspGlyThr 120
 Db 316 AGGGTTTGCCAGTTTCCCACTCCAGCTGCTGATCAGGGGAGCAGCTGAAGGAACT 375
 Qy 121 ValValLeuTrpAsnAlaGlnSerTrpLeuValArgCysGlySerValIleAspGly 140
 Db 376 GTGGTTTGTGGAATGACACAGTCATACAAATTATATGATGGTAGTAAAGATGGC 435
 Qy 141 SerLeuAlaAlaCysAlaPheSerProAsnGlySerPheValThrGlySerSerCys 160
 Db 436 TCCTTGGCGGCGATGTGATTTTCTCTATGGAAGCTTCTTGACAGCTGCTCTCATGT 495
 Qy 161 GlyAspLeuThrValITTPAspAspLysMetArgCysLeuHisSerGlyIleValHisAsp 180
 Db 496 GGTGATTTAACAGCTGTGGATGATAAATAGGTGTCTGCTGCTGATGAAAAGCAGATGAT 555
 Qy 181 LeuGlyIleThrCysCysAspPheSerSerGlnProValSerAspGlyGluGlnGlyLeu 200
 Db 556 CTGGAAATACCTCTCGATGATTTTCTTCAAGCAGAGTTCTGTATGAGAAACAAGGCTT 615
 Qy 201 GlnPhePheArgLeuAlaSerCysGlyGlnAspCysGlnValIleTrpIleValSer 220
 Db 616 CAGTTTTTGCACCTGGCATCATGTGTCCAGATTGCCAAGCAAAATTTGGATTGTTCT 675
 Qy 221 PheThrHisIleLeu----- 225
 Db 676 TTATCCATATCTTAGGTTTGAATTAATAATATAAGTACACTGAGTGGCACTGTGCT 735
 Qy 225 ----- 225
 Db 736 CTTGTCTGCTGTGTGCTTTTCCCATGATGGAGATGCTAGTCTAGGGTCAGTGAT 795
 Qy 225 ----- 225
 Db 796 AAGTCTGTCATGATATGATATCTAATACGAGAAATATCTTCAACATGACTCAGCAC 855
 Qy 225 ----- 225
 Db 856 ACCAGTATGTCACAACTGTGTGTTGACACTAATACCTTTTACTTGTACTGTTCA 915
 Qy 226 -----AlaArgArg 228
 Db 916 ATGACAAACAGTGAACATCTGGCAATTGACCTGGAACAATTGTCACAAAGGCGC 975
 Qy 229 ThrGlnIleGlnLeuValPheThrGlnAspTrpSerGluGlnValIleSerThrTrp 248
 Db 976 ACAGAAATCAGCTGAAGCAATTTTACGAGATTGGTCAGAGGAGATGTCTCAACATGG 1035
 Qy 249 LeuCysAlaGlnAspLeuLysAspLeuValGlyIlePheLysMetAsnIleAspGly 268
 Db 1036 CTGTGTGCACAAAGATTAAAGATCTTGTGTATTTTCAAGATGAATTAACATTGATGGA 1095
 Qy 269 LysGluLeuLeuAsnLeuThrLysGluSerIleuAlaAspAspLeuValIleGluSerLeu 288
 Db 1096 AAAAAGCTGTGATCTTCAAAAAGAAAGTGTGCTGATGATTTGAAAATTTGAATCTCTA 1155
 Qy 289 GlyLeuArgSerLysValIleuArgLysIleGluGlnLeuArgThrLysValIleSerLeu 308
 Db 1156 GGACTCGCTACTAAAGTCTGAGGAAATTTGAAGAGCTCAAGACCAAGTTAAATCCCTT 1215
 Qy 309 SerSerGlyIleProAspGluPheIleCysProIleThrArgGluLeuMetLysAspPro 328
 Db 1216 TCTTCAGGAATTCCTGATGAATTTATATGTCCAAATTAAGAACTTATGAAAGATCCG 1275

Qy 329 ValIleAlaSerAspGlyTrpSerTrpGluLysGluAlaMetGluAsnTrpIleSerLys 348
 Db 1276 GTCTATGCATCAATGATGCTTTCATATGAAAAGAAACATGAAATTTGATCAGCAA 1335
 Qy 349 LysLysArgThrSerProMetThrAsnLeuValLeuProSerAlaValLeuThrProAsn 368
 Db 1336 AAGAAAGCTACAAATGCCATGACAAATCTTGTCTTCTTCCAGCGGTAATTACACCAAT 1395
 Qy 369 ArgThrLeuLysMetAlaIleAsnArgTrpLeuGluThrHisGlnLys 384
 Db 1396 AGGACTCTGAAAATGGCCATCAATAGATGCTGGAGACACCAAAAG 1443
 RESULT 6
 ID AB224707 standard; cDNA; 1773 BP.
 AC AB224707;
 DT 07-APR-2003 (first entry)
 DE Human cell growth, differentiation and death protein CGDD-19 cDNA.
 XX 'CGDD-19'; cell growth; cell differentiation; cell death; human;
 KW cytostatic; antiarteriosclerotic; hepatotropic; antiinflammatory;
 KW antiparasitic; antianemic; ophthalmological; auditory; anticonvulsant;
 KW cerebroprotective; nootropic; neuroprotective; antiparkinsonian;
 KW neuroleptic; tranquilizer; immunosuppressive; anti-HIV; antiallergic;
 KW antidiabetic; antihypertoid; antidiabetic; dermatological; nephrotropic;
 KW antitumoric; antiarthritic; antidiabetic; antidiabetic; antidiabetic;
 KW antiparasitic; fungicide; antiparasitic; protozoicide; antihelminthic;
 KW antinfertility; gynecological; gene therapy; microarray; gene; ss.
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT CDS 60..1490
 FT /tag= a
 FT /product= "Human CGDD-19"
 XX
 XX WO200297032-A2.
 XX
 XX 05-DEC-2002.
 XX
 XX 05-APR-2002; 2002MO-US011152.
 XX
 XX 06-APR-2001; 2001US-0282110P.
 XX 11-APR-2001; 2001US-0283294P.
 XX 26-APR-2001; 2001US-0286820P.
 XX 27-APR-2001; 2001US-0287228P.
 XX 16-MAY-2001; 2001US-0291622P.
 XX 18-MAY-2001; 2001US-0291846P.
 XX 25-MAY-2001; 2001US-0293727P.
 XX 01-JUN-2001; 2001US-0295263P.
 XX 01-JUN-2001; 2001US-0295340P.
 XX 15-JUN-2002; 2002US-0349705P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 XX
 XX Aizimzai Y, Au-Young JK, Batra S, Baughn MR, Beach SD;
 PI Boroweky ML, Burford N, Ding L, Elliott VS, Emerling BW, Gandhi AR;
 PI Gietzen KJ, Griffin JA, Hatalia AJA, Honchell CD, Lal PG, Lee SY;
 PI Lu DM, Arvizu CS, Ramkumar J, Reddy R, Sanjanwala MM, Tang YT;
 PI Walla NK, Wang YS, Warren BA, Xu Y, Yang J, Yao MG, Yue H;
 PI Zebardjadian Y;
 XX
 XX WPI: 2003-140453/13.
 XX P-PSDB; ABP58348.
 XX
 XX Novel human proteins associated with cell growth, differentiation and
 PT death, useful for treating, diagnosing or preventing cancer,
 PT developmental, neurological, reproductive or autoimmune/inflammatory
 PT disorders.

XX Claim 5; Page 235-236; 238bp; English.

XX The present sequence is that of Incyte clone 7483131CB1 encoding human

CC CGDD-19, a novel protein associated with cell growth, differentiation and

CC death. A representative cDNA library for the polynucleotide is KIDNOT19

CC derived from kidney tissue. Structural features establish the encoded

CC protein as being associated with cell growth, differentiation and death,

CC and as showing sequence and structural similarity to human apoptotic

CC protease activating factor 1. The invention is based on novel human CGDD-1

CC (see AB24689-709), and to the use of these for the diagnosis, treatment

CC or prevention of cell proliferative disorders including cancer,

CC developmental disorders, neurological disorders, autoimmune disorders,

CC reproductive disorders, and disorders of the placenta, and in the

CC assessment of the effects of exogenous compounds on the activity and

CC expression of proteins and nucleic acids associated with cell growth,

CC differentiation and death. CGDD polynucleotides are also used in a

CC claimed microarray and in a claimed method of generating an expression

CC profile of a sample

XX

XX Sequence 1773 BP; 535 A; 334 C; 371 G; 533 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.35e-184	Length:	1773
Score:	1978.00	Matches:	382
Percent Similarity:	80.25%	Conservative:	0
Best Local Similarity:	80.25%	Mismatches:	2
Query Match:	96.63%	Indels:	92
DB:	8	Gaps:	1

US-10-077-111-13 (1-384) x AB224707 (1-1773)

QY	1	MetValIleuLeuIleIsthrLeuAlaAspHisGlyAspAspValAsnCysCysAlaPhe	20
DB	60	ATGCTGAATGATTCACACATTAGCTGATCATGTCACATGTCACATCTGCTTC	119
QY	21	SePheSerLeuLeuAlaThrCysSerLeuAspLysThrIleArgLeuIYrSerLeuArg	40
DB	120	TCTTTTCCCTCTTGCTACTTCTCTTGGACAAACATTCGCTGACTGTTACGT	179
QY	41	AspPheThrGluLeuProHisSerProLeuLeuPheHisThrIYrAlaValHisGlyCys	60
DB	180	GACTTACTGAACTGCACATTCCTCATTGAAGTTCAATCATGCTGTCACCTGTC	229
QY	61	CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrThrValIleu	80
DB	240	TGTTTCTCCCTTCAGACATATTTTGGCATCGTGTTCACACATGTCACCTGCT	239
QY	81	TrpAsnThrGluAsnGlyIleMetLeuAlaValMetGluIleProSerGlySerProVal	100
DB	300	TGGAAATACTGAAATGAGACAGATGCTGGCAGTATGGAACAGCCTAGTGCAAGCCCTG	359
QY	101	ArgValCysGlnInPheSerProAspSerThrCysLeuAlaSerGlyValAlaAspGlyThr	120
DB	360	AGGATTTGCGAGTTTCCCGACCTCCACGCTGTTGGCATCAGGGGCGACGTGATGAACT	419
QY	121	ValIleuLeuTrpAsnAlaGlnSerTrpLysLeuIYrArgCysGlySerValIYsAspGly	140
DB	420	GTGCTTTTGTGATGACACAGTCATACAAATTATATGATGTGTATGTTAAAGATGC	479
QY	141	SerLeuAlaAlaCysAlaPheSerProAsnGlySerPhePheValThrGlySerSerCys	160
DB	480	TCCTTGCGCGCATGTGCAATTTCTCTTAATGGAAGCTTCTTTGCACTGCTCCTCATGT	539
QY	161	GlyAspLeuThrValIYrPheAspAspLysMetArgCysLeuHisSerGlyValAlaHisAsp	180
DB	540	GGGATTTTAAACAATGTGGATGATTAATAAGAGTGTCTGATGTGAAAAAGCACTGAT	599
QY	181	LeuGlyIleThrCysAspAspPheSerSerGlnProValSerAspGlyGluIleuIYleu	200
DB	600	CTTGGATTATCCTGCTGCGCAATTTTCTTTCACAGCAGCTTCTGATGGAACAAGGCTT	659

QY	201	GlnPheAspArgLeuAlaSerCysGlyIleAspCysGlnValIleTrpIleValSer	220
DB	660	CAGTTTTTTCGACTGGCATCATGTGGTCAGATTCGCAAGTCAAATTTGGATTTGTTCT	719
QY	221	PheThrHisIleLeu-----	225
DB	720	TTTACCAATATCTTTAGCTTTTGAATTAATTAATAAGTACACTAGTGGGCACTGTGCT	779
QY	225	-----	225
DB	780	CCTGTTCTGCGCTTGCTTTTCCCATGATGGGACAGATGCTAGTCTCAGGGTCAGTGAAT	839
QY	225	-----	225
DB	840	AAGTCTGTATGATATGATGATCTAATACTGAAATATCTTACACATTGACTCAGCAC	899
QY	225	-----	225
DB	900	ACCAGGTATGTCAACATTGTGCTTTTGCACCTAATACCTTTTACTTGCTACTGTTCA	959
QY	226	-----AlaArgArg-----	228
DB	960	ATGGACAAACAGTGAACATCTGGCAATTGACCTGGAAACACTTTGCCAAGCAAGGAC	1019
QY	229	ThrGluHisGlnLeuLysGlnPheThrGluAspTrpSerGluGluValSerThrTrp	248
DB	1020	ACGAAACATCAGCTGAAAGCAATTTACCGAAGATGCTCAGAGAGAGATGTTCAACATCG	1079
QY	249	LeuCysAlaGlnAspLeuAspLeuValIleGlyIlePheLysMetAsnAsnIleAspGly	268
DB	1080	CTTTGTGCACAGATTTTAAAGATCTTGTGTTATTTTCAAGATGAATACATTGATGGA	1139
QY	269	LysGluLeuLeuAsnLeuThrLysGluSerLeuAlaAspAspLeuLysIleGluSerLeu	288
DB	1140	AAAGAACTTTGATCTTCAAAAGAACTCGGCTGATGATTTGAAATTTGAATCTGTA	1199
QY	289	GlyLeuAspSerLysValLeuArgLysIleGluGluLeuAspThrLysValLysSerLeu	308
DB	1200	GGACTGCTGATGAAAGTCTGAGAAATTTGAAGCTCAGGACCAAGCTTAAATCCCTT	1259
QY	309	SerSerGlyIleProAspGluPheIleCysProIleThrArgGluLeuMetLysAspPro	328
DB	1260	TCTTCAGAAATTCCTGATGAATTTATATGTCCAATACATACAGAACTTATGAAAGTCCG	1319
QY	329	ValIleAlaSerAspGlyIYrSerTrpGluLysGluAlaMetGluAsnTrpIleSerLys	348
DB	1320	GTCAATGCAATCAGATGCTATTCATATGAAAGAAAGCAATGGAATTTGATCAGCAAA	1379
QY	349	LysLysArgThrSerProMetThrAsnLeuValLeuProSerAlaValLeuThrProAsn	368
DB	1380	AAGAAACGTACAAAGTCCCATGACAAATCTTGTCTTCCCTTCAGCGTACTTACACCAAT	1439
QY	369	ArgThrLeuLysMetAlaIleAsnArgTrpLeuGluThrHisGlnLys	384
DB	1440	AGGACTCTGAAATGCGCATCAATGATGGCTGAGACACACCAAAAG	1487
RESULT 7			
ID	ADB63535	standard; cDNA; 1811 BP.	
AC	ADB63535;		
DT	04-DEC-2003	(first entry)	
XX	Human cDNA encoding clone TEST120264530.		
XX	Human; ss; gene; pharmaceutical; diagnostic; gene therapy;		
XX	tissue regeneration; cell regeneration; membrane protein;		
XX	signal transduction-related protein; transcription-related protein;		
XX	osteoporosis; neurological disease; cancer; tumour.		
XX	Homo sapiens.		
XX			

Key Location/Qualifiers
FT CDS 150..1580
FT /*tag= a
FT /product= "Clone TBSN10264530 protein"
XX
XX EPI308459-A2.
XX
XX 07-MAY-2003.
XX
XX 28-MAR-2002; 2002EP-00007401.
XX
XX 05-NOV-2001; 2001JP-00379298.
XX 25-JAN-2002; 2002US-00350978.
XX
XX (HELI-) HELIX RES INST.
XX (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
XX Iecgai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S,
XX Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I,
XX Seki N, Yoshikawa T, Otsuka M, Nagahari K, Maeuho Y;
XX
XX MPI: 2003-450961/43.
XX P-PSDB: ADB65505.
XX
XX New polynucleotides and polypeptides, useful for developing a diagnostic
XX marker or medicines for regulation of their expression and activity, or
XX as targets of gene therapy.
XX
XX
XX Claim 1; Page: 222pp; English.
XX
XX The invention discloses a polynucleotide comprising a sequence selected
XX from 1970 fully defined nucleotide sequences which encode novel
XX polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
XX or its partial peptide, an antibody binding to the polypeptide or peptide
XX of the polynucleotide, immunologically assaying the polypeptide or
XX peptide of the polynucleotide by contacting the polypeptide or peptide
XX with the antibody of the encoded protein, and observing the binding
XX between the two, a transformant carrying the polynucleotide in an
XX expressible manner and an antisense polynucleotide. The oligonucleotide
XX is useful as a primer for synthesizing the polynucleotide, or as a probe
XX for detecting the polynucleotide. The polynucleotides and encoded
XX proteins are useful as pharmaceutical agents and many disease-related
XX genes may be included in them, for developing a diagnostic marker or
XX medicines for regulation of their expression and activity, or as targets
XX of gene therapy. The genes are involved in tissue and/or cell
XX regeneration. Membrane proteins, signal transduction-related proteins,
XX transcription-related proteins, disease-related proteins and genes
XX encoding them can be used as indicators for diseases (e.g. osteoporosis,
XX neurological diseases, cancer, tumours. The cDNA may be used to regulate
XX the activity or expression of the encoded protein to treat diseases. The
XX sequence presented is a cDNA of the invention. Note: Some of the sequence
XX data for this patent is not represented in the printed specification, but
XX is based on sequence information supplied by the European Patent Office.
XX
XX Sequence 1811 BP; 520 A; 352 C; 399 G; 540 T; 0 U; 0 Other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 1.39e-184 Length: 1811
XX Score: 1978.00 Matches: 382
XX Percent Similarity: 80.25% Conservative: 0
XX Best Local Similarity: 80.25% Mismatches: 2
XX Query Match: 96.63% Indels: 92
XX DB: 10 Gaps: 1
XX
XX US-10-077-111-13 (1-384) x ADB63535 (1-1811)
XX
XX Qy 1 MetValIysLeuIleHisThrLeuAlaAspHisGlyAspAspValAsnCySValAlaPhe 20
XX Db 150 ATGATGAACCTGATTCACACATTCATGCTATGTCGATGCAATGCTGCTTC 209
XX Qy 21 SerPheSerLeuLeuAlaIleHisThrCysSerLeuAspIleHisGlyLeuIleSerLeu 40
XX Db 210 TCCTTTCCCTCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 269

Qy 41 AspPheThrGluLeuProHisSerProLeuIlePheHisThrTyraIaValHisCyS 60
Db 270 GACTTACTGAACTGACCACTTCTCCATTTGAAGTTTCACTATGCTGCTGCTGCTGCTG 329
Qy 61 CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrValLeu 80
Db 330 TGTTTCTCCCTTCAGAGACATATTTGGCATGCTGTTTCAACAGATGTTACACATGCTCTA 389
Qy 81 TrpAsnThrGluAsnGlyGlnMetLeuAlaValMetGlnIleProSerGlySerProVal 100
Db 390 TGGATATCTGAAATGAGACAGATGCTGCGATGAGTGAACGCTTACTGAGCCCTGCTG 449
Qy 101 ArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyAlaIleAspGlyThr 120
Db 450 AGGTTTGGCAGATTTTCCCAAGCTCCAGACTCCAGTGTGGCATCAAGGGGACGCTGATGAACT 509
Qy 121 ValValLeuThrPheAlaGlnSerTyrlsLeuIleTyraGlySerValIleAspGly 140
Db 510 GTGGTTTGTGGAATGACACAGTCATCAAAATATATGATGTGTGTTAAAGATGCG 569
Qy 141 SerLeuAlaIleCysAlaPheSerProAsnGlySerPhePheValThrGlySerSerCys 160
Db 570 TCCTTGGCGGACATGTGATTTTCTCTTAATGAAAGCTTGTGTCACTGCTCCTCATGT 629
Qy 161 GlyAspLeuThrValTrpAspAspIleMetArgCysLeuHisSerGlyIleValAlaHisAsp 180
Db 630 GGTGATTTAAACATGTGGATGATTAATAGATGTCTGATAGTGAAGCAACATGAT 689
Qy 181 LeuGlyIleThrCysCysAspPheSerSerGlnProValSerAspGlyGlnGlyLeu 200
Db 690 CTGGAATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 749
Qy 201 GlnPhePheArgLeuAlaSerCysGlyGlnAspCysGlnValIleIleIleValSer 220
Db 750 CAGTTTTTGTGACGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 809
Qy 221 PheThrHisIleLeu----- 225
Db 810 TTTTACCATATCTTAGGTTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 869
Qy 225 ----- 225
Db 870 CTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 929
Qy 225 ----- 225
Db 930 AAGTCTGTCATAGTATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 989
Qy 225 ----- 225
Db 990 ACCAGTATGTCACAACTTGCTGCTTGGACCTTAATACCTTTTACTGCTACTGCTTCA 1049
Qy 226 -----AlaArgArg 228
Db 1050 ATGACAAACAAACAGTGAACATCTGGCATTTGACCTGGAACACTTGGCAAGCAAGAGAC 1109
Qy 229 ThrGlnHisGlnLeuIleGlnPheThrGlnAspTrpSerGlnGlnValIleSerThrTrp 248
Db 1110 ACAGAAATCATGAGTGAAGCAATTTTACCGAAGATTTGACAGAGGAGATGCTCAACATGG 1169
Qy 249 LeuCysValIleGlnAspLeuIleAspLeuValGlyIlePheIleHisMetAsnAsnIleAspGly 268
Db 1170 CTTTGTGACAAAGATTTTAAAGATCTTGTGTGATTTTCAAGATTAATTAATTAATTAATGA 1229
Qy 269 LysGlnLeuLeuAsnLeuThrIleGlnSerLeuAlaAspAspLeuIleGlnSerLeu 288
Db 1230 AAAGAACTGTTGATATCTTAACAAAGAAAGCTGGCTGATATTTGAAATTTGATATCTTA 1289
Qy 289 GlyLeuArgSerIleValIleLeuArgIleGlnGlnLeuArgThrIleValIleSerLeu 308
Db 1290 GGACTGCTGATGAAGCTGAGGAAATGAAAGAGCTCAGACCAAGGTTTAAATCCCTT 1349

QY 309 SerSerGlyIleProAspGluPheIleCysProIleThrArgGluLeuWetLysAspPro 328
DB 1350 TCTTCAGAGATTCCTGATGAAATTAATGTCCAATACTAGAGAACTTATGAAAGATCCG 1409
QY 329 ValIleAlaSerAspGlyTyrSerTyrGluLysGluAlaMetGluMetPrlIleSerLys 348
DB 1410 GTATGCGATCAATGCGATATCATATGAAAGAGCAATGGAATTTGATCAGCAAA 1469
QY 349 LysLysArgThrSerProMetThrAsnLeuValLeuProSerAlaValLeuThrProAsn 368
DB 1470 AAGAAAGTACAAATGCCATGACAAATCTGTTCTTCTCAGGGGTACTTACACCAAT 1529
QY 369 ArgThrLeuLysMetAlaIleAsnArgTyrPLeuGluThrHisGluLys 384
DB 1530 AGGACTCTGAATAATGGCCATCAATAGATGGCTGGAGACACCAAAAG 1577
RESULT 8
ADB47507
ID ADB47507 standard; cDNA; 1817 BP.
XX
AC ADB47507;
XX
XX 04-DEC-2003 (first entry)
DE Human cDNA upregulated in dendritic cells SEQ ID NO 207.
XX
XX aa; gene; human; dendritic cells; high throughput; cancer;
KM infectious disease; autoimmune disease; allergy;
KM graft versus host disease; vaccine enhancing; gene therapy.
XX
OS Homo sapiens.
XX
XX US2003134283-A1.
XX
PD 17-JUL-2003.
XX
PF 03-OCT-2001; 2001US-00971392.
XX
PR 03-OCT-2000; 2000US-0237652P.
XX
PA (PETE/) PETERSON D P.
PA (PEAR/) PEARSON C I.
PA (COCK/) COCKS B G.
XX
XX Peterson DP, Pearson CI, Cocks BG;
XX
XX WPI; 2003-662509/62.
XX
PT New combination comprises cDNAs that are differentially expressed in
PT dendritic cells useful for preparing a composition for diagnosing or
PT treating cancer, infectious disease, autoimmunity, allergy or graft
PT versus host disease.
XX
PS Claim 1, SEQ ID NO 207; 28bp; English.
XX
XX The invention relates to a combination comprising cDNAs that are
XX differentially expressed in dendritic cells (DC). Also included is a high
XX throughput method for detecting differential expression of one or more
XX cDNAs in a sample containing nucleic acids. The combination is useful for
XX preparing a composition for diagnosing, treating and monitoring the
XX treatment of cancer, infectious disease, autoimmunity, allergy or graft
XX versus host disease, or for enhancing a vaccine. The present sequence
XX represents a human cDNA upregulated in dendritic cells. Note: The
XX sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format directly from USPTO
XX at seqdata.uspto.gov/sequence.html?docID=20030134283.
XX
SQ Sequence 1817 BP; 519 A; 373 C; 383 G; 542 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1,396-184 Length: 1817
Score: 1978.00 Matches: 382
Percent Similarity: 80.25% Conservative: 0

Best Local Similarity: 80.25% Mismatches: 2
Query Match: 96.63% Indels: 92
DB: 10 Gaps: 1
US-10-077-111-13 (1-384) x ADB47507 (1-1817)
QY 1 MetValLysLeuLleHisThrLeuAlaAspHisGlyAspAspValAsnCysGlyAlaPhe 20
DB 148 ATGGTGAACCTGAACTGATTCACACATTAGCTGATCGTGAAGAGATGTCAACTGCTGCTTC 207
QY 21 SerPheSerLeuLeuAlaThrCysSerLeuAspLysThrIleArgLeuTyrSerLeuArg 40
DB 208 TCTTTTCCCTCTTGACCTCTTGCTCTTGACCAAAACAATTCGCTGACTGCTTACCT 267
QY 41 AspPheThrGluLeuProHisSerProLeuLysPheHisThrLysAlaValLysCys 60
DB 268 GACTTACTGAACCTGACACATTCATTCATGAAAGTTTCACTACCTAGTCTCCACCTGCTG 327
QY 61 CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrValLeu 80
DB 328 TGTTCCTCCCTTCAGACATATTTTGGCATGTGTTCAACAGATGTACACTGTCTTA 387
QY 81 TrpAsnThrGluAsnGlyGlnMetLeuAlaValMetGluGlnProSerGlySerProVal 100
DB 388 TGGAAATACGAAATGAGACAGATGCTGGCAGATGAGAGAGCTTAATGACAGCCCTGTG 447
QY 101 ArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyAlaAlaAspGlyThr 120
DB 448 AGGGTTTGGCAAGTTTCCCAAGCTCCACGTGTGTTGGCATCAGGGGCACTGATGGAAT 507
QY 121 ValValLeuTrpAsnAlaGlnSerTyrLysLeuTyrArgCysGlySerValLysAspGly 140
DB 508 GTGGTTTGTGGAATCAGATCATCAATTAATTAATGATGTGTGTAAAGATGCG 567
QY 141 SerLeuAlaAlaCysAlaPheSerProAsnGlySerPheValThrGlySerSerCys 160
DB 568 TCTTTGCGGCAATGATTTCTCTTAATGAGAGCTTCTTGTCACCTGCTCTCATATG 627
QY 161 GlyAspLeuThrValTrpAspAspLysMetCysValLeuHisSerGluLysAlaHisAsp 180
DB 628 GGTGATTTAAACAGTGTGGGATGATTAATGAGGTGTCTGCATATGAAAGACACATGAT 687
QY 181 LeuGlyIleThrCysAspPheSerSerGlnProValSerAspGlyGluGlnGlyLeu 200
DB 688 CTGGAAATTAACCTGCTGCATTTTCTTCACAGCCAGTTCGATGGAACAACAGGCTCTT 747
QY 201 GlnPhePheArgLeuAlaSerCysGlyGlnAspCysGlnValLysIleTrpIleValSer 220
DB 748 CAGTTTTTTCGACTGGCATCATGTGTGTCAGGATTTGCCAAGTAATTTGGATGTTC 807
QY 221 PheThrHisIleLeu----- 225
DB 808 TTTTACCATATCTTAGGTTTGAATTAATAATTAAGTACAGTGGGCACTGTGCT 867
QY 225 ----- 225
DB 868 CTTGTTGCGCTTGCTCTTTTCCATGATGGCAGATGCTAGTCTCAGGGTCAAGTGAAT 927
QY 225 ----- 225
DB 928 AAGTCTGTATAGTATATGATACATAATGAGAAATATATCTTACACATGACTGACAG 987
QY 225 ----- 225
DB 988 ACCAGTATGTCACAACTGTGTGCTTGGACCTTAATCCCTTTTACTGCTACTGTGTC 1047
QY 226 -----Ala-Arg- 228
DB 1048 ATGAGCAAAACAGTGAACATCTGGCAATTGACCTGGAACAATTGCGCAAGCAAGAGAC 1107
QY 229 ThrGluHisGlnLeuLysGlnPheThrGluAspTrpSerGlyGluValLysThrTrp 248
DB 1108 ACAGAACATCAGCTGAAGCAATTTTACCGAAGATGTGTCAAGAGAGATGTCTCAACATGG 1167


```
QY 201 GlnPhePheArgLeuAlaSerCysGlyGlnAspCysGlnValIleThrIleValSer 220
    |||
DB 764 CAGTTTTTTCGACTGGCATCATGTGTCAGGATTCGCAAGTCAAAATTGGATTGTTCT 823
QY 221 PheThrHisIleLeu----- 225
    |||
DB 824 TTATCCCATATCTTAGGTTTGAATTAAATATTAAGTCACTGAGTGGCACTGTGCT 883
QY 225 ----- 225
DB 884 CCGTTCTGGCTTGCTTTTCCCATGATGGCGAGATGCTAGCTCAGGGTCAGTGAT 943
QY 225 ----- 225
DB 944 AAGCTGTGATAGTATATGATTAATACTGAGATAATATCTGCACATTGACTCAGAC 1003
QY 225 ----- 225
DB 1004 ACCAGGTATGTCAACAATTGTGCTTTGCACTTAATACCTTTTACTTGCTACTGTTCA 1063
QY 226 -----AlaArg 228
    |||
DB 1064 ATGGACAACAGTGAACATCTGGCAATTGACCTGGAACACTTTGCCAAGCAGAGAC 1123
QY 229 ThrGluHisGlnLeuLysGlnPheThrGluAspTrpSerGluGluValIleSerThrTrp 248
    |||
DB 1124 ACAGAACATCAGCTGGAAGCAATTATCCGAGATTTGTCAGAGAGAGATGTCTCAACATCG 1183
QY 249 LeuCyAlaGlnAspLeuLysAspLeuValGlyIlePheLysMetAsnIleAspGly 268
    |||
DB 1184 CTTTGTCACAGATTTTAAAGATCTGTGTATTTTCAAGATGAATTAACATTGATGGA 1243
QY 269 LysGluLeuLeuAsnLeuThrLysGluSerLeuAlaAspAspLeuLysIleGluSerLeu 288
    |||
DB 1244 AAGAACTGTGATCTTACAAAGAAAGTCTGGCTGATGATTTGAAATTTGATCTCTA 1303
QY 289 GlyLeuArgSerLysValLeuArgLysIleGluGlnLeuArgThrLysValLysSerLeu 308
    |||
DB 1304 GGACTGGGTAGTAAAGTCTGAGGAAATTAAGAGCTCAGACCAAGGTTAAATCCCTT 1363
QY 309 SerSerGlyIleProAspGluPheIleCysProIleThrArgGluLeuMetLysAspPro 328
    |||
DB 1364 TCTTCAGGAATCTCTGATGAATTAATATGTCACATTAATCTGAGAACTTAAGAAAGATCCG 1423
QY 329 ValIleAlaSerAspGlyTrpSerTrpGluLysGluLysIleMetGluAsnTrpIleSerLys 348
    |||
DB 1424 GTCATCGCATCAAGATGCTATTCATATGAAAGAAAGCAATGAAATTTGATCAGCAAA 1483
QY 349 LysLysArgThrSerProMetThrAsnLeuValLeuProSerAlaValLeuThrProAsn 368
    |||
DB 1484 AAGAAACGTACAAAGTCCCATGACAAATCTTCTCTTCAAGCGGTACTTACACCAAT 1543
QY 369 ArgThrLeuLysMetAlaIleAsnArgTrpLeuGluThrHisGlnLys 384
    |||
DB 1544 AGGACTCTGAAATATGGCATCAATAGATGGCTGAGACACACCAAAAG 1591

RESULT 10
ADQ9098
ID ADQ9098 standard, cDNA; 1844 BP.
XX
AC ADQ9098;
XX
DT 23-SEP-2004 (first entry)
XX
DE DNA encoding human GPCR-like protein seqid 768.
XX
KW ophthalmological; immunomodulatory; cytosstatic; antiatherosclerotic;
XX antiidiabetic; GPCR-like protein; ophthalmic disorder;
KW neurological disorder; immunological disorder; nephritic disorder;
KW hormonal dysfunction; cancer; atherosclerosis; diabetes;
KW molecular weight marker; food supplement; human; ss.
OS Homo sapiens.
```

```
XX
PN US659662-B1.
XX
PD 27-MAY-2003.
XX
PF 19-JUL-2000; 2000US-00620312.
XX
PR 21-JAN-2000; 2000US-00488725.
XX
PR 25-APR-2000; 2000US-00552317.
XX
PA (HYSE-) HYSER INC.
XX
PI Tang YT, Zhou P, Dymnac RT;
XX
DR WPI; 2001-442255/47.
XX
PT New G-protein-coupled receptor-like polypeptides and polynucleotides.
PT useful for treating diseases of ophthalmic, neurological, immunological
PT and nephritic systems and hormonal dysfunction, cancer, atherosclerosis
PT and diabetes.
XX
PS Example 2; SEQ ID NO 768; 92pp; English.
XX
CC The invention describes an isolated polynucleotide (I) comprising a fully
CC defined (SI) of 749, 3188, 2484, 1169, 2936, 1467, 5773, 5714, 4041,
CC 1372, 3996, 3945, 2735, 1788, 585, 1782, 927, 5714 or 2282 nucleotides as
CC given in the specification, its translated or protein coding portion, its
CC extracellular portion or its active domain. The GPCR-like polypeptides
CC and polynucleotides are useful for the treatment of diseases of
CC ophthalmic, neurological, immunological and nephritic systems. They may
CC also be used to treat hormonal dysfunction, cancer, atherosclerosis and
CC diabetes. The antibodies are useful for detecting or quantitating the
CC polypeptide in tissue. The polypeptides can also be used as molecular
CC weight markers and as a food supplement. This sequence represents a human
CC polynucleotide of the invention.
XX
SQ Sequence 1844 BP; 535 A; 380 C; 389 G; 540 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 1,42e-184 Length: 1844
Score: 1978.00 Matches: 382
Percent Similarity: 80.25% Conservative: 0
Best Local Similarity: 80.25% Mismatches: 2
Query Match: 96.63% Indels: 92
DB: 5 Gaps: 1

US-10-077-111-13 (1-384) x ADQ9098 (1-1844)
QY 1 MetValLysLeuLleHisThrLeuAlaAspHisGlyAspAspValAsnCysCyAlaPhe 20
    |||
DB 164 ATGGTGAACCTGATTCACACATTAGCTGATCAATGAGTGAAGATGTCMACTGCTGTGCTTC 223
QY 21 SerPheSerLeuLeuAlaThrCysSerLeuAspLysThrIleArgLeuTrpSerLeuArg 40
    |||
DB 224 TCCTTTTCCTCTTGCTGCTATCTGCTCTTGCAAAACAATTCGCTGACTGTTACGT 283
QY 41 AspPheThrGluLeuProHisSerProLeuLysPheHisThrTrpAlaValHisCysCys 60
    |||
DB 284 GACTTACTGAACTGCAATCTCCATTGAAGTTTCAATCACTATGCTCCACTGCTGC 343
QY 61 CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrThrValLeu 80
    |||
DB 344 TGTTCCTCCCTTCAGAGCATATTTTGGCATGTGTGTTCAACAGATGATGCCACTGTCTA 403
QY 81 TrpAsnThrGluAsnGlyGlnMetLeuAlaValMetGluGlnProSerGlySerProVal 100
    |||
DB 404 TGGAAATACGAAATATGACAGATGCTGCGAGATGGAACACCTTAAGTGCAGCCCTGTG 463
QY 101 ArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyValAlaAspGlyThr 120
    |||
DB 464 AGGGTTTGCACAGTTTCCCAAGCTCCACGCTGTTGGCATCAGGGGACGACTGATGAAC 523
QY 121 ValValLeuTrpAsnAlaGlnSerTrpLysLeuTrpArgCysGlySerValLysAspGly 140
```

```

Db      524 GTGGTTTGTGGATGACAGTCAATATTAATATGATGTGGTAAAGATGCG 583
Qy      141 SerLeu1Aa1aCySa1aPheSerProsnGlySerPheValThrGlySerSerCys 160
Db      584 TCCTGGGGGATGGATGCTTTCTCTAATGGAAGCTTCTTGTCAGCGGCTCCATGAT 643
Qy      161 GlyAspLeuThrVal1TRPAspAspLysMetArgCysLeuHisSerGluLysAlaHisAsp 180
Db      644 GGTATTTTAACAGTGTGGATGATTAATAAGTGTCTGCATAGTGAAGAAAAGCATGAT 703
Qy      181 LeuGly1IleThrCysCysAspPheSerSerGlnProVal1SerAspGlyGluGlnGlyLeu 200
Db      704 CTTCGAATTAACCTGCTGGATTTTCTTCACAGCCAGTTCTGCATGAGAACAAAGTCTT 763
Qy      201 GlnPhePheArgLeuAla1SerCysGlyGlnAspCysGlnValLys1IleTrp1IleValSer 220
Db      764 CAGTTTTCGATGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 823
Qy      221 PheThrHis1IleLeu----- 225
Db      824 TTTACCCATATCTTAGTGTGTTGAATTAATAATTAATAAGTACATGAGTGGCATGTGCT 883
Qy      225 ----- 225
Db      884 CCTGTTCTGGCTGTGCTTTTCCCATGATGGGAGATGATGATGATGATGATGATGAT 943
Qy      225 ----- 225
Db      944 AAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1003
Qy      225 ----- 225
Db      1004 ACCAGATATGTCACAACTTGCTTTTGACCTAATACCTTTTACTTGTCTACTGTTCA 1063
Qy      226 ----- 228
Db      1064 ATGACAAACAGTGAACATCTGGCAATTTGACCTGAAACACTTTGGCCAGCAAGAGAC 1123
Qy      229 ThrGluHisGlnLeuLysGlnPheThrGluAspTrpSerGluGluVal1SerThrTrp 248
Db      1124 ACAGAACCTACAGCTGAGCAATTTTACCCAAATTTGCTGAGGAGGATGCTTCAACATGG 1183
Qy      249 LeuCysAlaGlnAspLeuLysAspLeuValGly1IlePheLysMetAsn1IleAspGly 268
Db      1184 CTCTTGTCACAAATTTTAAAGATCTTGTTGATTTTCAAGATGAATTAACATGATGGA 1243
Qy      269 LysGluLeuLeuAsnLeuThrLysGluSerLeuAlaAspAspLeuLys1IleGluSerLeu 288
Db      1244 AAAGACCTGTTGAATCTTACAAAGAAAGTCTGCTGATGATTTGAAATTTGAATCTCTA 1303
Qy      289 GlyLeuAspSerLysValLeuAspGly1IleGluGluLeuAspGlyThrLysValLysSerLeu 308
Db      1304 GGACTGCTAGTAAAGTCTGAGGAAATTTGAAAGCTCAGAGCCAAAGTTAAATCCCTT 1363
Qy      309 SerSerGly1IleProAspGluPhe1IleCysPro1IleThrArgGluLeuMetLysAspPro 328
Db      1364 TCTTCAGGAATTCCTGATGAATTTATATGTCAAATTAAGTAACTTAAGAACTTATAAAGATCCG 1423
Qy      329 Val1IleAlaSerAspGlyTrpSerTrpGluLysGluAlaMetGluAsnTrp1IleSerLys 348
Db      1424 GTCAATCGCATCAAGATGCTATTTCAATATGAAAGAAAGAAAGAAATTTGATCAGCAAA 1483
Qy      349 LysAlaValArgThrSerProMetThrAsnLeuValLeuProSerAlaValLeuThrProsn 368
Db      1484 AAGAAAGCTACAAATGCCATCAAAATCTTCTCTTCTCAGCGGTACTTACACCAAT 1543
Qy      369 ArgThrLeuLysMetAla1IleAsnArgTrpLeuGluThrHisGlnLys 384
Db      1544 AGGACTCTGAATAATGGCCATCAATATGATGCTGGAGACACACCAAAAG 1591

```

```

ID      ADB48858 standard; cDNA; 1844 BP.
XX
AC      ADB48858;
XX
DT      04-DEC-2003 (first entry)
XX
DE      Novel human cDNA SEQ ID NO 768.
XX
KW      ss; cancer; neurodegenerative disease; human.
XX
OS      Homo sapiens.
XX
PN      US2003104529-A1.
XX
PD      05-JUN-2003.
XX
PF      04-JAN-2002; 2002US-00037270.
XX
PR      21-JAN-2000; 2000US-00488725.
PR      25-APR-2000; 2000US-00552317.
PR      19-JUL-2000; 2000US-00620312.
XX
PA      (ZHOU/) ZHOU P.
PA      (TANG/) TANG Y T.
PA      (LIUC/) LIU C.
PA      (ASUN/) ASUNDI V.
PA      (DRMA/) DRMANAC R T.
PI      Zhou P, Tang YF, Liu C, Asundi V, Drmanac RT;
PI      WPI; 2003-678194/64.
XX
PT      New polynucleotide, useful for treating diseases e.g., cancer or
PT      neurodegenerative diseases.
XX
PS      Claim 1; SEQ ID NO 768; 99p; English.
XX
CC      The invention relates to a polynucleotide comprising a sequence given in
CC      the specification, or its mature protein-coding portion, or its
CC      complement. The polynucleotide is useful for treating diseases e.g.,
CC      cancer or neurodegenerative diseases and many others listed in the
CC      specification. The present sequence represents a novel human cDNA. Note:
CC      The sequence data for this patent did not form part of the printed
CC      specification but was obtained in electronic format directly from USPTO
CC      at Seqdata.uspto.gov/sequence.html?docID=20030104529.
XX
SQ      Sequence 1844 BP; 535 A; 380 C; 389 G; 540 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 1,42e-184 Length: 1844
Score: 1978.00 Matches: 382
Percent Similarity: 80.25% Conservative: 0
Best Local Similarity: 80.25% Mismatches: 2
Query Match: 96.63% Indels: 92
DB: 9 Gaps: 1
US-10-077-111-13 (1-384) x ADB48858 (1-1844)
Qy      1 MetValLysLeu1IleHisThrLeuAlaAspHisGlyAspAspValAsnCysCysAlaPhe 20
Db      164 ATGGTGAATCTGATTCACATTAAGTGAATGATGATGATGATGATGATGATGATGATGAT 223
Qy      21 SerPheSerLeuLeuAlaThrCysSerLeuAspLysThr1IleArgLeuTrpSerLeuArg 40
Db      224 TCCCTTTCCCTCTTGGCTACTTCTCTCTGACAAACAAATTCGCTGATCTGTTAAGT 283
Qy      41 AspPheThrGluLeuProHisSerProLeuLysPheHisThrTrpAlaValHisCysCys 60
Db      284 GACTTATCTAAGTACGACATCTTCAATTAAGTTCAATCAATGATGATGATGATGATGAT 343
Qy      61 CysPheSerProSerGlyHis1IleLeuAlaSerCysSerThrAspGlyThrThrValLeu 80
Db      344 TGTTTCTCCCTTCAGGACATATTTTGGCATGCTGTTCAACATGATGATGATGATGATGAT 403

```

QY 81 TTPAsnThrGluangnglygImeLleuAlValMeGluGlnProSerGlySerProVal 100
 DB 404 TGGAAATACGAAATAGGACAGATGCTGGCAGTGTGAAACAGGCTAGTGGCAGCCCTGTG 463
 QY 101 ArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyValAlaAspGlyThr 120
 DB 464 AGGGTTTGGCAGTTTCCAGACTCCACGTTGTCATCAGGGGGCAGCTGATGAACT 523
 QY 121 ValValLeuThrPheAlaGlnSerTyrLygLeuThrArgCysGlySerValLygAspGly 140
 DB 524 GTGGTTTGTGGAATGACAGTCACTACAAATTATATGATGTGTTAAAGATGGC 563
 QY 141 SerLeuAlaAlaCysAlaPheSerProAspGlySerPhePheValThrGlySerSerCys 160
 DB 584 TCCCTGGCGGCACTGTCATTTTCTCTTAATGGAAGCTCTTTGTCACTGGCTCCTCATGT 643
 QY 161 GlyAspLeuThrValTTPAspAspLysMetArgCysLeuHisSerGlyValAlaHisAsp 180
 DB 644 GGTGATTTTAAACAGTGTGGATGATAAATAGAGTGTCTGATAGTGAAGAACACATGAT 703
 QY 181 LeuGlyLeuThrCysCysAspPheSerSerGlnProAlaSerAspGlyGluGlnGlyLeu 200
 DB 704 CTGGAAATTAACCTGCTGCGATTTTCTTCAACGCCAGCTTCTGATGAGAACAAAGCTCT 763
 QY 201 GlnPhePheArgLeuAlaSerCysGlyGlnAspCysGlnValLygIleTPIIleValSer 220
 DB 764 CAGTTTTTTCGACCTGGCATCATGTGTCAGAGTTGCCAAAGTCAAAATTGGATTGTTCT 823
 QY 221 PheThrHisIleLeu----- 225
 DB 824 TTAAACCATATCTTAGGTTTGAATTAAATATATAAGTACACTAGAGTGGCACTGTGCT 883
 QY 225 ----- 225
 DB 884 CCTGTCTGCGCTGTGCTTTTCCCATGATGGCAGATGCTAGTCTCAGGGTCACTGGAT 943
 QY 225 ----- 225
 DB 944 AAGCTGTCAATAGTATGATACTAATACAGAGAATTACTTCAACATGACTCAGCAC 1003
 QY 225 ----- 225
 DB 1004 ACCAGTATGTCACAACTGTGCTTTGCACTAATACCTTTTACTTGCTACTGTCTCA 1063
 QY 226 -----AlaArgArg 228
 DB 1064 ATGACAAACAAAGTGAACATCTGGCAATTGACCTGAAACACTTTGCCAAGCAAGAGC 1123
 QY 229 ThrGluHisGlnLeuLysGlnPheThrGluAspTyrSerGluGluValValSerThrTyr 248
 DB 1124 ACAGAACTACGCTGGAAGCAATTTACCGAAGATTGGTCAGAGAGAGATGTCTCAACATGG 1183
 QY 249 LeuCysAlaGlnAspLeuLysAspLeuValGlyIlePheLysMetAsnIleAspGly 268
 DB 1184 CTGTGTGCACAAAGTTTAAAGATCTGTGGTATTTTCAAGATGAATTAACATTGATGGA 1243
 QY 269 LygGluLeuLeuAsnLeuThrLygGluSerLeuAlaAspAspLeuLysIleGluSerLeu 288
 DB 1244 AAGAACTGTGATCTTACAAAGAAAGCTGGCTCATGATTTGAAATTTGAATCTCTTA 1303
 QY 289 GlnLeuArgSerLysValLeuArgLygIleGluGluLeuArgThrLysValLysSerLeu 308
 DB 1304 GGACTGGTGAAGTAAAGTCTGAGGAAATTTGAAGAGTCTGAGACCAAGTTAAATCCCTT 1363
 QY 309 SerSerGlyIleProAspGluPheIleCysProIleThrArgGluLeuMetLysAspPro 328
 DB 1364 TCTTCAGAGAAATCTCTGATGAATTTATATGTCCATTACTAGAGAACTTAAGAAAGATCCG 1423
 QY 329 ValIleAlaSerAspGlyTyrSerTyrGluLygGlnAlaMetGluAsnTPIIleSerLys 348
 DB 1424 GTCATCCGATCAGATGCTATTCTATATGAAAGAAAGCAATGCAAAATTTGATCAGCAAA 1483

QY 349 LysLygArgThrSerProMetThrAsnLeuValLeuProSerAlaValLeuThrProAsn 368
 DB 1484 AAGAAACGTACAAAGTCCCATGACCAATCTTGTCTTCTTCAGCGGTACTTACCAACAAAT 1543
 QY 369 ArgThrLeuLysMetValIleAsnArgTPIIleAsnGluThrHisGlnLys 384
 DB 1544 AGGACTCTGAATAATGGCCATCAATGATGCTGGAGACACACCAAAAG 1591
 RESULT 12
 ID AAD45076 standard; DNA; 1908 BP.
 AC AAD45076;
 DT 27-DEC-2002 (first entry)
 XX
 DE Human RET16.3 splice variant DNA.
 XX
 KW Human; RET16; intracellular signal; inflammation-related disease; asthma;
 KW rheumatoid arthritis; psoriasis; multiple sclerosis; neuroprotective;
 KW transplant rejection; chronic obstructive pulmonary disease; TNF-alpha;
 KW inflammatory bowel disease; Crohn's disease; immunosuppressive; cancer;
 KW inacute respiratory distress syndrome; cardiac; ulcerative colitis;
 KW autoimmune disease; cystic fibrosis; gene therapy; cytostatic; neoplasm;
 KW cellular migration disorder; cell proliferation disorder; calcification;
 KW hyperinsulinaemia; diabetes type 2; systemic lupus erythematosus; tumour;
 KW cardiovascular disease; Wegener's granulomatosis; atherosclerosis;
 KW thalassemia; vasotropic; gene; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT 136..164
 FT CDS /tag= a
 FT /product= "Human RET16.3 splice variant protein"
 XX
 MO200266494-A2.
 XX
 PD 29-AUG-2002.
 XX
 PF 15-FEB-2002; 2002MO-US005162.
 XX
 PR 16-FEB-2001; 2001US-0269366P.
 PR 29-MAY-2001; 2001US-0294181P.
 XX
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX
 PI Todderud CG, Finger JN, Rillema J;
 XX
 DR WPI; 2002-682760/73.
 XX
 DR P-PSDB; AAE28168.
 XX
 PT New human, mouse or rat RET16 genes and proteins, involved in
 PT intracellular signaling cascade, useful for in gene therapy, particularly
 PT for treating e.g. ischemia, cystic fibrosis, autoimmune disease, cancers,
 PT tumors or neoplasms.
 PT
 XX
 PS Claim 1; Page 168-169; 175pp; English.
 XX
 CC The invention relates to human, mouse or rat RET16 genes and proteins,
 CC involved in intracellular signaling cascade. The RET16 protein or
 CC polynucleotide is useful for treating an inflammation-related disease or
 CC disorder, e.g. rheumatoid arthritis, juvenile arthritis, psoriasis, or
 CC asthma, ischaemia-reperfusion, multiple sclerosis, rejection of organ or
 CC tissue transplants, chronic obstructive pulmonary disease, inflammatory
 CC bowel disease, Crohn's disease, ulcerative colitis, inacute respiratory
 CC distress syndrome, systemic lupus erythematosus, autoimmune disease,
 CC cystic fibrosis, cancers, tumours or neoplasms. This inflammation-related
 CC disease or disorder also includes disorders associated with aberrant
 CC activation of the TNF-alpha pathway, disorders associated with aberrant
 CC cellular migration, proliferation, metastasis, juvenile idiopathic
 CC arthritis, haematogenous metastases of tumour cells, hyperinsulinaemia,
 CC diabetes type 2, atherosclerosis, cardiovascular disease, colon cancer,

PI Mao Y, Xie Y;
 XX WPI; 2001-537038/60.
 DR P-PSDB; AAG78660.
 XX
 PT New polypeptide-beta-transducin 41 for treating e.g. cancer and HIV
 PT infection.
 XX
 PS Claim 6; Page 24-25 (Disclosure); 32pp; Chinese.
 CC The present invention provides the protein and coding sequences of beta-
 CC transducin 41. The sequences can be used in the treatment of cancer and
 CC HIV infection. The present sequence is the coding sequence of the
 CC invention.
 XX
 SO Sequence 1623 BP; 485 A; 302 C; 335 G; 501 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,966-172	Length:	1623
Score:	1854.00	Matches:	364
Percent Similarity:	79.18%	Conservative:	1
Best Local Similarity:	78.96%	Mismatches:	2
Query Match:	90.57%	Indels:	94
DB:	4	Gaps:	1

US-10-077-111-13 (1-384) x AA164914 (1-1623)

QY	18	CysAlaPheSerPheSerLeuLeuAlaThrCysSerLeuAspLeuThrIleArgLeuTyr	37
DB	3	TGAGCTTCTCTCTTTTCCCTCTTGCTACTTCTCTTGGACAAACAAATTCGCTGTAC	62
QY	38	SerLeuArgAspPheThrGluLeuProHisSerProLeuLysPheHisThrTyrAlaVal	57
DB	63	TCGTTAGAGACTTACTTGAACGCCAATTCCTCATTAAGTTTCTACTAGCTGCTC	122
QY	58	HisCysCysCysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThr	77
DB	123	CACGTGCTGCTGTTCTCCCTTCAGGACATATTTTGGCATGCTGTTCCAAACAGATGGTACC	182
QY	78	ThrValLeuThrPheAnthrGluAsnGlyGlnMetLeuAlaValMetGluGlnProSerGly	97
DB	183	ACTGCTCTATGAAATCTGAAATGACAGATGCTGACAGTGAAGAACCCCTAGTGGC	242
QY	98	SerProValArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyAlaAla	117
DB	243	AGCCCTGTGAGGCTTCCAGTTTCCCAAGCTCCACGTGTTGGCATCAGGGGCACT	302
QY	118	AspGlyThrValValLeuThrPheAlaGlnSerTyrIleLeuTyrArgCysGlySerVal	137
DB	303	GATGGAACCTGCTGTTTGTGGAATGACAGTCAATACAAATTAATAGATGGTAGTGT	362
QY	138	LysAspGlySerLeuAlaAlaCysAlaPheSerProAsnGlySerPhePheValThrGly	157
DB	363	AAAGATGGCTCTTGCGGCGCATGTCATTTCTCTTAATGAAAGCTTCTTGTCACGGC	422
QY	158	SerSerCysGlyAspLeuThrValThrAspAspLysMetArgCysLeuHisSerGlyLys	177
DB	423	TCCCTCATGTGTGATTTAAACAGTGTGGAGATTAATAGAGTGTCTGCATAGGAAAA	482
QY	178	AlaHisAspLeuGlyIleThrCysAspPheSerSerGlnProValSerAspGlyGlu	197
DB	483	GCACATGATCTTGGAATATTAATCTGCTGCATTTTCTTCAACAGCCAGTTTCTGATGGAA	542
QY	198	GlnGlyLeuGlnPhePheArgLeuAlaSerCysGlyGlnAspCysGlnValLysIleTyr	217
DB	543	CAAGGTTTCAAGTTTTCAGCTGGCATCAATGTGTGAGGATTCGCAAGTCAAAATTTGG	602
QY	218	IleValSerPheThrHisIleLeu-----	225
DB	603	ATTGTTTCTTTTACCCCATATCTTAGCTTTGGAATTAATAATAAAGTACAGTGAAGG	662
QY	225	-----	225

DB	663	CACGTGCTCTGTTCTGCTGCTGCTTTTCCCATGATGGCAGATGCTAGTCTCAGG	722
QY	225	-----	225
DB	723	TCAGTGATTAAGTCTGTCTATAGTATATGATAGTAACTAGAGATTAATCTTGCACATNG	782
QY	225	-----	225
DB	783	ACTCAGACACAGATGATGTCAAACTGTGCTTTTGGACACTAATACCTTTTACTTGCT	842
QY	225	-----	225
DB	843	ACTGTTCAATGACAAACAGTGAACATCTGGCAATTTGACCTGGAACACTTGGCAA	902
QY	226	AlaArgArgThrGluHisGlnLeuLysGlnPheThrGluAspTrpSerGluValVal	245
DB	903	GCAAGGCGCAGACATACATGCTGAAGCAATTTACGAAAGATTGGTCAGAGAGAGATGC	962
QY	246	SerThrTrpLeuCysAlaGlnAspLeuLysAspLeuValGlyIlePheLysMetAsnAsn	265
DB	963	TCAACATGGCTTTGTGCACAAATTTAAAGATCTTGTGTGATTTTCAAGATGAATTAAC	1022
QY	266	IleAspGlyLysGluLeuLeuLeuThrLysGlnSerLeuAlaAspAspLeuLysIle	285
DB	1023	ATTGATGAAAGAACTGTTGAATCTTACAAAGAAAGCTGCTGATGATTTGAAATTT	1082
QY	286	GluSerLeuGlyLeuArgSerLysValLeuArgLysIleGluGluLeuArgThrLysVal	305
DB	1083	GAATCTTAGACTGCTGATGTAAGTCTGAGGAAATTTGAAGAGCTCAGACCAAGTT	1142
QY	306	LysSerLeuSerSerGlyIleProAspGluPheIleCysProIleThrArgGluLeuMet	325
DB	1143	AAATCCCTTTCTCAGAAATTCCTGATGAATTAATATGTCCAATAACTGAGAAACTTANG	1202
QY	326	LysAspProValIleLeuAspArgLysTyrSerTyrGluLysGluAlaMetGluLeuTyr	345
DB	1203	AAAGATCCGCTCATGCTGATCAATGATGCTATCAATGAAAGAAAGCAATGAAATTTGG	1262
QY	346	IleSerLysLysIleLysArg-ThrSerProMetThrAsnLeuValLeuProSerAlaValLe	365
DB	1263	ATCAGCAAAAGAAACGGTACAAAGTCCCATGCAAAATCTTGTCTTCTTCAAGCGGCT	1322
QY	365	u-ThrProAsnArgThrLeuLysMetAlaIleAsnArgTrpLeuGluThrHisGlnLys	384
DB	1323	TTACACCAATATGAACTTGAAATGCGCATCAATGATGCTGAGGACACACCAAAAG	1381

RESULT 14
 AA160662/c
 ID AA160662 standard; cDNA; 1826 BP.
 XX
 AC AA160662;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 4651.
 XX
 KW Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; hemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemoclastic;
 KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
 KW leukemia; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-US034263.
 XX
 PR 23-DEC-1999; 99US-00471275.
 PR 21-JAN-2000; 2000US-00488725.

PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX
XX (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,
PI Wang J, Wang Z, Weinman T, Xu C, Xue AD, Yang Y, Zhang J, Zhao QA,
PI Zhou P, Goodrich R, Dirmacac RT;
XX
XX WPI: 2001-442253/47.
DR P-PSDB; AAM41506.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders such
XX as central nervous system injuries.
XX
XX Claim 1; SEQ ID NO 4651; 10078bp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and the
XX encoded polypeptides (AAM38642-AA42213) with neurotropic,
XX immunosuppressant and cytostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localized neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilization of the activities such as: Immune system suppression,
XX Activin/inhibin activity, chemotactic/chemokinetic activity, hemostatic
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX assays for receptor activity, arthritis and inflammation, leukemias and
XX C.N.S disorders. Note: The sequence data for this patent did not form
XX part of the printed specification
XX
XX Sequence 1826 BP; 535 A; 382 C; 380 G; 529 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 2.6e-172 Length: 1826
XX Score: 1853.50 Matches: 370
XX Percent Similarity: 77.34% Conservative: 2
XX Best Local Similarity: 76.92% Mismatches: 12
XX Query Match: 90.55% Indels: 98
XX DB: Gaps: 4
XX
XX US-10-077-111-13 (1-384) x AAI60662 (1-1826)
XX
XX 1 MetValIysLeuIleHisThrLeuAlaAspHisGlyAspAspValaSerCysAlaPhe 20
XX 1701 ATGTGTGAACGATTCACACATTCATGCTGATGTCAGTGCATGCTGTCCTTC 1642
XX
XX 21 SerPheSerLeuLeuAlaThrCysSerLeuAspLysThrIleArgLeuIleYSerLeuArg 40
XX 1441 TCCCTTTCCCTCTGGCTACTGCTCTCTTCGACAAACAATTCGCTGATCTGTTACT 1582
XX
XX 41 AspPheThrGluLeuProHisSerProLeuLysPheHisThrIleAlaValHisCysCys 60
XX 1581 GACTTTACTGAACGACCATTCCTCCATGTAAGTTTCATACCTATGCTGCACCTGCGC 1522
XX
XX 61 CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrThValleu 80
XX 1521 TGTTCCTCCCTTCAGGACATATTTTGGCATCGTTTCACAGATGATGACCATGCTCTA 1462
XX
XX 81 TrpAsnThrGluAsnGlyGluMetLeuAlaValMetGluGluProSerGlySerProVal 100
XX 1461 TGGATTAATCTGAATGACAGATGCTGACATGATGACAGACCTAGTGCGACCTCTGTG 1402
XX
XX 101 ArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyAlaAlaAspGlyThr 120
XX 1401 AGGGTTTCCAGTTTTCCTCCAGACTCCAGTGTTTGGCATCAGGGGCGAGCTGATGGAAC 1342

QY 121 ValValIleuTrpAsnAlaGlnSerIleYsLeuIleYArgCysGlySerValLysAspGly 140
DB 1341 GTGGTTTGTGGATGACAGCTCATCAAAATTAATATGATGTGTGTGTTAAAGATGCG 1282
QY 141 SerLeuAlaAlaCysAlaPheSerProAsnGlySerPhePheValThrGlySerSerCys 160
DB 1281 TCCCTGGCGGCGATGTGATTTTCTCCAAAGAAAGCTTCTTGTCTACTGCTCCCTCAATG 1222
QY 161 GlyAspLeuThrValIleTrpAspAspLysMetArgCysLeuHisSerGlyLysValHisAsp 180
DB 1221 GGTGATTTAAGATGAGGATGATTAATATGATGATGATGATGATGATGATGATGATGAT 1162
QY 181 LeuGlyIleThrCysCysAspPheSerSerSerGlnProValSerAspGlyGluGlnGlyLeu 200
DB 1161 CTGGATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1102
QY 201 GlnPhePheArgLeuAlaSerCysGlyGlnAspCysGlnValIleYsIleTrpIleValSer 220
DB 1101 CAGTTTTCGACTGCGCATGATGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1042
QY 221 PheThrHisIleLeu----- 225
DB 1041 TTTAACCATATCTTAGGTTTGAATTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 982
QY 225 ----- 225
DB 981 CCGTTTGGCTGTGTGCTTTTCCCATGATGGGACAGATGCTAGTCTCAGGTCAGTGGAT 922
QY 225 ----- 225
DB 921 AAGTCTGATATGATATATGATATGATATGATATGATATGATATGATATGATATGATATGAT 862
QY 225 ----- 225
DB 861 ACCAGTATGTCACAACTTGCTGCTTGGACCTTAATACCTTTTAATCTGCTACTGTTCA 802
QY 226 -----AlaArgArg 228
DB 801 ATGACAAACAAACAGTGAACATCTGCAATTTGACCTGGAACACCTTGCCAAAGAGAGC 742
QY 229 ThrGluHisGlnLeuLysGlnPheThrGluAspTrpSerGluGluValIleSerThrTrp 248
DB 741 ACAGAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 682
QY 249 LeuCysAlaGlnAspLeuLysAspLeuValGlyIlePheLysMetCysAsnIleAspGly 268
DB 681 CTTTGTGACACAAATTTAAAGATCTTGTGTGATATTTTCAAGATGATTAATCATGATGATGA 622
QY 269 LysGluLeuLeuAsnLeuThrLysGluSerLeuAlaAspAspLeuLysIleGluSerLeu 288
DB 621 AAAGAACTGTGATCTTAACAAAGAAAGCTGCTGATATGATATGATATGATATGATATGAT 562
QY 289 GlyLeuArgSerLysValLeuArgLysIleGluGluLeuArgThrLysValLysSerLeu 308
DB 561 GGACTGCTGATGAAGCTGAGGAAATTTGAAGAGCTCAGACCAAGATTAAATCCTT 502
QY 309 SerSerGlyIleProAspGluPheIleCysProIleThrArgGluLeuMetLysAspPro 328
DB 501 TCTTCAGAAATTCCTGATGATATTAATGATGATGATGATGATGATGATGATGATGATGAT 442
QY 329 ValIleAlaSerAspGlyLysSerIleGluGluAlaMetGluLeuMetLysSerLys 348
DB 441 GTATGTCATCAAGATGCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 383
QY 349 LysLeuArg-----ThrSerProMetThrAsn-----LeuValLeuProSerAla-----Val 364
DB 382 AAAAGAAACCGTAACAGTCCCTCATGACAAATCCTTGTCTTCCCTTCACGCGATAC 323
QY 365 LeuThrProAsnArgThrLeuLysMet-AlaIleAsnArgTrpLeuGluThrHisGlnLys 384
DB 322 TTACACCCAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 263

OY 384 s 384
Db 262 A 262
RESULT 15
ADM19385
ID ADM19385 standard, cDNA, 1686 BP.
XX
AC ADM19385;
XX
DT 20-MAY-2004 (first entry)
XX
DE Novel human channel/transporter gene #182.
XX
XX ds; gene; immunosuppressive; antiarthritic; antiheumatic;
KM antiproliferative; cytosolic; cardiac; vasoprotective;
KM neurotropic; neuroprotective; antibacterial; virucide; fungicide;
KM optalmological; gene therapy; channel/transporter protein;
KM rheumatoid arthritis; neoplasia; cardiac arrest; cerebrovascular disorder;
KM cerebral ischemia; angiogenesis; nervous system disorder;
KM Alzheimer's disease; ocular disorder; corneal infection; wound healing;
KM epithelial cell proliferation; skin aging; sunburn; transplantation;
KM chemotaxis; food additive.
XX
OS Homo sapiens.
XX
PN WO200154472-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001307.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184654P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-019076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215115P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218280P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226682P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0228924P.
PR 30-AUG-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.

PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-024060P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.

PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249219P.
 PR 17-NOV-2000; 2000US-0249224P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249279P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX WPI; 2001-476159/51.
 DR P-PSDB; ADM19864.
 XX
 PT Isolated nucleic acid molecule encoding a channel/transporter protein is
 PT used in preventing, treating or ameliorating a medical condition.
 PS Claim 1; SEQ ID NO 192; 809pp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule encoding a
 CC channel/transporter protein or sequences at least 95% identical to a
 CC these. The nucleic acids and proteins encoded by them are used to
 CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
 CC in diagnosing a pathological condition or susceptibility to a
 CC pathological condition. The antibodies to the proteins can also be used
 CC in alleviating symptoms associated with the disorders and in diagnostic
 CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays
 CC (ELISA). Disorders which are diagnosed or treated include autoimmune
 CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
 CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
 CC arrest, cerebrovascular disorders e.g. cerebral ischemia, angiogenesis,
 CC nervous system disorders e.g. Alzheimer's disease, infections caused by
 CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection.
 CC The polypeptides can also be used to aid wound healing and epithelial
 CC cell proliferation, to prevent skin aging due to sunburn, to maintain
 CC organs before transplantation, for supporting cell culture of primary
 CC tissues, to regenerate tissues and in chemotaxis. The polypeptides can
 CC also be used as a food additive or preservative to increase or decrease
 CC storage capabilities. This sequence corresponds to a gene of the
 CC invention.
 XX
 SQ Sequence 1686 BP; 485 A; 348 C; 353 G; 500 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 9 2e-159 Length: 1686
 Score: 1715.50 Matches: 335
 Percent Similarity: 70.38% Conservative: 0
 Best Local Similarity: 70.38% Mismatches: 2
 Query Match: 83.81% Indels: 139
 DB: 5 Gaps: 2
 US-10-077-111-13 (1-384) x ADM19385 (1-1686)
 QY 1 MetValIysLeuIleHisThrLeuAlaAspHisGlyAspAspValaAsnCySCysAlaPhe 20
 DB 150 ATGGTGAACATGATTCACACATTAAGTCTGATCATGTGCAATGTCATGCTGCGCTTC 209

QY SerPheSerLeuLeuAlaThrCysSerLeuAspLysThrIleArgLeuIYrSerLeuArg 40
 DB 210 TCCTTTTCCTCTTGGGATCTTGTCTCTTGGACAAACATTCGCTGATCTGTTACGT 269
 QY AspPheThrGluLeuProHisSerProLeuLysPheHisThrIYrAlaValHisCysCys 60
 DB 270 GACTTACTGACATGCACATTCCTCCATTAAGTTTCATACCTATGCTGTCCACTGCTGC 329
 QY CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrIYrPheGlyThrThrValLeu 80
 DB 330 TGTTCCTCCCTTCAGACATATTTTGGCATCGTGTTCACAAATGATGTACCTACCTCTTA 389
 QY TrpAsnThrGluAsnGlyGlnMetLeuAlaValMetGluGlnProSerGlySerProVal 100
 DB 390 TGGAAATCTGAAAATGACAGATCTGCGAGTGTGATGGAACAGCTTACTGACAGCCCTGTG 449
 QY ArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyAlaIaAspGlyThr 120
 DB 450 AGGTTTGGCCAGTTTCCAGACTCCACGTTTGGCATCAGGGGAGCTGATGAACT 509
 QY ValValLeuThrPheAlaGlnSerIYrLysLeuIYrArgCysGlySerValIYrAspGly 140
 DB 510 GTGTTTGTGGATGACACAGTCATACAAATTAATATGATGTGTACTGTAAAGATGCG 569
 QY SerLeuAlaIaCysAlaPheSerProAsnGlySerPhePheValThrGlySerSerCys 160
 DB 570 TCCTTGCGGCAATGTCATTTCTCTTAATGAAGCTTCTTGTCTACTGCTCTCTCATGT 629
 QY GlyAspLeuThrValTrpAspAspLysMetArgCysLeuHisSerGlyLysAlaHisAsp 180
 DB 630 GGTGATTTTACAGTGTGGATGATTAATAGGTGTCTGATGTGAATAAGCACAATGAT 689
 QY LeuGlyIleThrCysCysAspPheSerSerGlnProValSerAspGlyGlnGlyLeu 200
 DB 690 CTGGAAATTAACCTGCTGCGATTTTCTTTCACACCGAGTTCTGATGGAACAAGGCTT 749
 QY GlnPhePheArgLeuAlaSerCysGlyGlnAspCysGlyValIYrIlePheValSer 220
 DB 750 CAGTTTGTGACCTGCATCATGTGTGTCAGAGTTGCCAATGCAAAATTTGGATTGTCT 809
 QY PheThrHisIleLeu----- 225
 DB 810 TTTAACCATATCTTAGCTTTTAATTAATTAATTAAGTACACTGAGTGGCACTGTGCT 869
 QY 225 ----- 225
 DB 870 CTTGTTCTGGCTTGTGCTTTTTCATGATGGCAGATGCTAGTCTCAGGGTCAATGAT 929
 QY 225 ----- 225
 DB 930 AAGTCTGTCAATGATTAATGATTAATTAATGATTAATGATTAATGATTAATGATTAATGAT 989
 QY 225 ----- 225
 DB 990 ACCAAGTATGTCAACACTGTGCTTTGACACTTAATCCTTTTACTTGTCTACTGTTCA 1049
 QY 226 -----AlaArgArg 228
 DB 1050 ATGACAAACACATGTAACATCTGCAATTTGACCTGGAACACTTTGGCCAAAGAGAGC 1109
 QY ThrGluHisGlnLeuLysGlnPheThrGluAspTrpSerGluGluValIaSerThrTrp 248
 DB 1110 ACAGAAATCATGCTGAAGCAATTTACCGAAGATTTGTCAGAGAGAGATGTCCTCAACATGG 1169
 QY LeuCysAlaGlnAspLeuLysAspLeuValIYrIlePheLysMetAsnAsnIleAspGly 268
 DB 1170 CTTTGTGACAAAGATTTTAAAGATCTTGTGTGATTTTCAAGATGAATTAATGATGATGA 1229
 QY LysGluLeuLeuAsnLeuThrIYrGlySerLeuAlaAspAspLeuLysIleGlySerLeu 288
 DB 1230 AAAGAACTGTGATCTTAACAAAGAAAGTCTGGCTGATGATTTGAAAATTT----- 1280

```

Qy      289 GlyLeuArgSerIysValIeuArgLysIleGluGluLeuArgThrIysSerLeu 308
Db      1280 ----- 1280
Qy      309 SerSerGlyIleProAspGluPheIleCysProIleThrArgGluLeuMetLysAspPro 328
Db      1280 ----- 1280
Qy      329 ValIleAlaSerAspGlyTyrSerTyrGluLysGluAlaMetGluAsnTrpIleSerLys 348
Db      1281 -----GATGGCTATTTCATATGAAAAGAGAGCAATGGAATAATTGATCAGCAAA 1328
Qy      349 LysLysArgThrSerProMetThrAsnLeuValIeuProSerAlaValIeuThrProAsn 368
Db      1329 AAGAAACGTACAGTCCCATGACAAATCTGTCTTCCTTCAGCGGTACTTACACCAAAAT 1388
Qy      369 ArgThrLeuLysMetAlaIleAsnArgTrpLeuGluThrHisGlnLys 384
Db      1389 AGGACTCTGAAATGGCCATCAATAGTGGCTGGAGACACCAAAAG 1436

```

Search completed: February 5, 2005, 12:07:55
 Job time : 723 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 5, 2005, 11:39:07 ; Search time 4204 Seconds
(without alignments)
3476.850 Million cell updates/sec

Title: US-10-077-111-13
Perfect score: 2047
Sequence: 1 MVXIHITLADHGDDVNCAP.....LTPNRLKMAINWLETHOK 384

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 segs, 19032134700 residues
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USFTO.spool/US10077111/runat_04022005_121053_6816/app.query.fasta_1.583
-DB=EST -QMT=faetap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blousum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pico -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRR=US10077111 @CGN 1.1 4352 @runat_04022005_121053_6816 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -MAIN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:
1: gb_est1:*
2: gb_est2:*
3: gb_hlc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_ges1:*
9: gb_ges2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1977	93.6	1770	3	CR607851 full-length
2	1714.5	86.8	1600	3	CR619104 full-length
3	1617	79.0	2027	3	AK011391 Mus muscu
4	1205	58.9	889	1	AL553333
5	1203	58.8	941	1	AL533510
6	1154	56.4	1010	4	BM543484
7	1043.5	50.4	858	4	BI603184
8	1031.5	49.1	792	4	BI821143
9	1005		728	7	CN793033

10	975.5	47.7	910	5	B0895162	B0895162	AGENCOURT
11	943.5	46.1	838	4	BI107478	BI107478	AGENCOURT
12	907	44.3	839	7	CK306752	CK306752	AGENCOURT
13	883.5	43.2	808	7	CK792250	CK792250	AGENCOURT
14	880	43.0	499	5	BX282729	BX282729	AGENCOURT
15	880	43.0	686	4	BI759505	BI759505	AGENCOURT
16	879	42.9	727	7	CK364577	CK364577	AGENCOURT
17	873	42.6	505	1	AI189142	AI189142	AGENCOURT
18	872	42.6	807	1	AL040518	AL040518	AGENCOURT
19	856.5	41.8	817	5	B0961462	B0961462	AGENCOURT
20	854	41.7	718	5	B0356767	B0356767	AGENCOURT
21	834	40.7	652	2	BB653743	BB653743	AGENCOURT
22	821	40.1	669	1	AJ726222	AJ726222	AGENCOURT
23	801	39.1	869	5	BU148588	BU148588	AGENCOURT
24	797	38.9	760	5	BU273077	BU273077	AGENCOURT
25	758.5	37.0	965	4	BG769613	BG769613	AGENCOURT
26	758	37.0	683	6	CD215933	CD215933	AGENCOURT
27	747	36.5	919	5	BQ713478	BQ713478	AGENCOURT
28	746	36.4	791	7	CK603751	CK603751	AGENCOURT
29	739	36.1	682	6	BY710675	BY710675	AGENCOURT
30	725	35.4	768	7	CK846814	CK846814	AGENCOURT
31	719	35.1	559	4	BI681850	BI681850	AGENCOURT
32	715	34.9	967	1	AL533462	AL533462	AGENCOURT
33	705	34.4	797	5	BG143666	BG143666	AGENCOURT
34	700	34.2	764	4	BG619844	BG619844	AGENCOURT
35	698.5	34.1	767	7	CN526819	CN526819	AGENCOURT
36	691	33.8	723	9	CE487278	CE487278	AGENCOURT
37	690	33.7	632	4	BJ624489	BJ624489	AGENCOURT
38	677	33.1	626	4	BG562985	BG562985	AGENCOURT
39	673.5	32.9	447	1	AA118718	AA118718	AGENCOURT
40	668.5	32.7	1042	4	CO567478	CO567478	AGENCOURT
41	662.5	32.4	829	7	CO567478	CO567478	AGENCOURT
42	662	32.3	463	7	CN407189	CN407189	AGENCOURT
43	662	32.3	897	5	BQ964199	BQ964199	AGENCOURT
44	657	32.1	1010	3	BC039804	BC039804	AGENCOURT
45	648	31.7	699	1	AJ726210	AJ726210	AGENCOURT

ALIGNMENTS

RESULT 1
CR607851
LOCUS full-length cDNA clone CS0DN0047015 of Adult brain of Homo sapiens (human).
DEFINITION CR607851 1770 bp mRNA linear HTC 21-JUL-2004
ACCESSION CR607851 GI:50488658
VERSION CR607851.1
KEYWORDS HTC; CNS/IT CDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1770)
AUTHORS Li, W.B., Gruber, C., Jesse, J., and Polayes, D.

TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished

REMARK Contact: Feng Liang Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

REFERENCE 2 (bases 1 to 1770)
AUTHORS Genoscope.

TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail: seqref@genoscope.cns.fr

COMMENT

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime ends enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.

FEATURES

source 1..1770
/organism="Homo sapiens"


```

/mol_type="mrna"
/db_xref="taxon:9606"
/clone="CSOD10751B08"
/tissue_type="Placenta
/plasmid="pCMTSPORT_6"

```

[illegible]

US-10-077-111-13 (1-384) X CR619104 (1-1600)

Oy		MetValysLeuIleHisTrpLeuAlaAspHisGlyAspArgValAsnCysGlyAlaPhe	20
Db	112	ATGTGGAAACTGATTACACATTAGCTGCATGTCGACGATGTAACCTGCTGCCCTTC	177
Oy	21	SerPheSerLeuLeuAlaThrCysSerLeuAspLysThrIleArgLeuLysSerLeuArg	40
Db	172	TCCTTTTCCTTGCGTCACTTCTCCTTGACCAAAACAATCGCCTGAACGTTAAGT	233
Oy	41	AspPheThrGluLeuProHissSerProLeuLysPheHisThrTyrAlaValHisCysCys	60
Db	232	GACTTAACTGAACTGGCACATTTCCATTGAAGTTTCATACCTAATGCTGCACGCTGC	293
Oy	61	CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrThrValLeu	80
Db	292	TGTTTTCCCCCTTCAGACACATAATTTGGCATGCTGTTCACAGATGATACCATCTCTTA	353
Oy	81	TyrAsnThrGluAsnGlyInMetLeuAlaValMetGluGlnProSerGlySerProVal	100
Db	352	TGGAAATACTGAAAATGACAGATGCTGCGACATGATGAAACAGCTTAGCGAGCCCTGTG	413
Oy	101	ArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyValAlaAspGlyThr	120
Db	412	AGGGTTTCCAGTTTTCGCCAGCTCAGGTGTTTGGCATCAGGGCAGCTGATGAACT	473
Oy	121	ValValLeuTrpAsnAlaGlnSerTyrLysLeuTyrArgCysGlySerValLysAspGly	140
Db	472	GTGGTTTTTGGAATGCACAGTCATCAAAATTATATAGTGTGATAGTTTAAAGATGC	533
Oy	141	SerLeuAlaAlaCysAlaPheSerProAsnGlySerPhePheValThrGlySerSerCys	160
Db	532	TCCTTGGCGGCAAGTCATTTCTCTTAATGAAAGCTTTCTTGTCACGCGCTCCCAAGT	593
Oy	161	GlyAspLeuThrValTrpAspAspLysMetArgCysLeuHisSerGluLysAlaHisAsp	180
Db	592	GGTGATTTAACAGTGTGGATGTATATAATGAGGTGTCCGATATGTAAGAAGACATGAT	653
Oy	181	LeuGlyIleThrCysCysAspPheSerSerGlnProValSerAspGlyGluGlnGlyLeu	200
Db	652	CTTGGAATTACCTGCTGCCATTTTCTTCACAGCCAGTTTGCATGGAGAAACAAGCTTT	713
Oy	201	GlnPhePheArgLeuAlaSerCysGlyGlnAspCysGlnValLysIleTrpIleValSer	220
Db	712	CAGTTTTTTCAGCTGGCATCATGTGTCCAGATGCGCAAGTCAAAATTTGATGTGTTCT	773
Oy	221	PheThrHisIleLeu-----	222
Db	772	TTTATCCATATCTTAGGTTTGAATTAAATATAAAGTACACTGATGGGCACTGTCT	833
Oy	225	-----	225
Db	832	CCTGTTCTGCGCTTGCTTTTCCATGATGGCAGATGCTAGTCTCAGGGTCAGTGAT	893
Oy	225	-----	225
Db	892	AAGTCTGTCAATAGTATATGATACTAAATCTGAAATATATCTTCACACATTGACTCAGAC	953

[illegible]

RESULT 3	AK011391	LOCUS	DEFINITION	2027 bp	mRNA	Linear	HTC 03-APR-2004
REFERENCE	AK011391		Mus musculus 10 days embryo whole body cDNA, RIKEN full-length enriched library, clone:2610014f08 product:hypothetical SAM domain (Serile alpha motif)/Modified RING finger domain/G-protein beta WD-40 repeats containing protein, full insert sequence.				
ACCESSION	AK011391						
VERSION	AK011391.1		GI:12847483				
KEYWORDS	HTC; CAP trapper.						
SOURCE	Mus musculus (house mouse)						
ORGANISM	Mus musculus						
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.						
AUTHORS	1						
TITLE	Carninci, P. and Hayashizaki, Y						
JOURNAL	High-efficiency full-length cDNA cloning						
MEDLINE	Weth. Enzymol. 303, 19-43 (1995)						
PUBMED	99279253						
REFERENCE	10349636						
AUTHORS	2						
TITLE	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.						
JOURNAL	Normalization and subtraction of cap-trapper-selected cDNAs to						
MEDLINE	prepare full-length cDNA libraries for rapid discovery of new genes						
PUBMED	Genome Res. 10 (10), 1617-1630 (2000)						
REFERENCE	11042159						
AUTHORS	3						
TITLE	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, T., Nishi, K., Kitanai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakauechi, S., Ikegami, T., Kaishiwagi, K.,						


```

Db      823 ATTACCGGTCTTAGGCTTTGAATTAATAAAGACACATTAAGTGGCGACTGGCC 882
Qy      222 -----
Db      883 CCTGTTCGGCTGTGCTTTTTCATCATGATGGAAGAATGCTGCATGGGGCTAGTGAT 942
Qy      223 -----
Db      943 AAATCTGTATCATATACATGATGCGCCCTCAGAGTGTGTACACAGCGTCACTAGCAGAT 1002
Qy      229 ThrluHts-----HisileuHts-----
Db      1003 ACCAGGATGTATGACATTTGCGCTTGCACCAACTCTTACTTGTACTGTTCA 1062
Qy      231 -----
Db      1063 ATGACAAAGACATGACATTTGGCAGTTTACCTGGAACACCTTGCAGGAAGCATG 1122
Qy      232 -----GluLeuYsglnPheThrluAspTyrSerGluGluValSerThrlu 249
Db      1123 AACGACCGCTGAAACATTTTCAATGATGTCAGAGGAGATGTCCTCGTGTGGTT 1182
Qy      250 CysAlaGlnAspLeuYsAspLeuValGlyIlePheYsMetAsnAsnIleAspGlyLys 269
Db      1183 CGTCTCAAGGCTTGAAAGACCTCGTGTATTTTCAGGGCAACAACATGATGGGAAA 1242
Qy      270 GluLeuLeuAsnLeuThrluGluSerLeuAlaAspAspLeuYsIleGluSerLeuGly 289
Db      1243 GAATCTATTGCACTTCAACAAAGAAAGTGTGCTGTGATTTGAAATGAAATCTCTAGGG 1302
Qy      290 LeuYsSerLysValLeuYsIleGluGluLeuYsThrluYsValYsSerLeuSer 309
Db      1303 CTGGCGAGCAAAAGCTCTGAGGATTAAGAGAGCTCAGGGCCAGATGATCTCTCTCT 1362
Qy      310 SerGlyIleProAspGluPheIleCysProIleThrluGluLeuMetLysAspProVal 329
Db      1363 TCCGGAATCCCTGACGAGTTCATCTGCCAATAACAGAACTCATGAAGAACCCGCTC 1422
Qy      330 IleIleAspAspGlyYrSerYrGluYsGluValMetGluAsnTyrIleSerLysLys 349
Db      1423 ATCCATCAATGCTCTCTCTACAGAGAACCAAGCAACTGATCACAAGAG 1482
Qy      350 LysArgThrSerProMetThrAsnLeuValLeuProSerAlaValIleThrluProAsnArg 369
Db      1483 AAGCTACGAGCCCAATGCAAAATTTGGCTCTCCCTTCACTGATGACCCCAAGAG 1542
Qy      370 ThrluYsMetAlaIleAsnArgTyrLeuGluThrluHisGlnLys 384
Db      1543 AACTGAAGATGGCCATCAACCGATGGCTGAGACGACAGAGAG 1587

RESULT 4
AL553333 889 bp mRNA linear EST 30-MAR-2004
LOCUS AL553333 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS001075YB08 5-PRIME, mRNA sequence.
ACCESSION AL553333
VERSION AL553333.3 GI:45858102
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (base 1 to 889)
L1,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:31275147.
Contact: Genoscope - Centre National de Sequencage
Genoscope - Centre National de Sequencage
2 rue Gaëton Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: gscrf@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned

```

```

FEATURES
source
1..889
/organism="Homo sapiens"
/mol_type="mRNA"
/di_xref="taxon:9606"
/clone="CS001075YB08"
/rna_type="PLACENTA COT 25-NORMALIZED"
/clone_id="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Alignment Scores:
Pred. No.: 5.48e-112 Length: 889
Score: 1205.00 Matches: 222
Percent Similarity: 99.12% Conservative: 2
Best Local Similarity: 98.23% Mismatches: 1
Query Match: 58.87% Indels: 1
DB: 1 Gaps: 0

US-10-077-111-13 (1-384) x AL553333 (1-889)
Qy      1 MetValYsLeuIleHisThrLeuAlaAspHisGlyAspAspValAenCysCysAlaPhe 20
Db      112 ATGTGTGAACATGATTCACACATTAAGTGAATGATGATGATGATGATGATGATGATG 171
Qy      21 SerPheSerLeuLeuAlaThrCysSerLeuAspLysThrIleArgLeuYrSerLeuArg 40
Db      172 TCTTTTCCCTGCTGCTACTTGTGCTCTTGGACAAACAAATTCGCTGACTGATGATGATG 231
Qy      41 AspPheThrGluLeuProHisSerProLeuLysPheHisThrYrAlaValHisCysCys 60
Db      232 GACTTTACTAATCAATGACCAATTCCTCAATTAAGTTTATATCAATGCTGCTGCTGCTG 291
Qy      61 CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrValLeu 80
Db      292 TGTTCCTCCCTTCAGGACATATTTTGGCATCGTTTCAACAGATGATGATGATGATGATG 351
Qy      81 TrpAsnThrGluAsnGlyGlnMetLeuAlaValMetGluGlnProSerGlySerProVal 100
Db      352 TGGAAATCTGAATAATGACAGATGCTGCAATGATGATGATGATGATGATGATGATGATG 411
Qy      101 ArgValCysGln-PheSerProAspSerThrCysLeuAlaSerGlyValAlaAspGlyThr 120
Db      412 AGGGTTTGGCAGTTTTCCTCCAGCATCCAGCTGTTTGGCATGAGGGCAGCTGATGAGAC 471
Qy      120 ValValLeuThrAsnAlaGlnSerYrLysLeuYrArgCysGlySerValYsAspGly 140
Db      472 TGTGTTTGTGGATGACAGATGATCAATTAATTAATTAATTAATTAATTAATTAATTAATG 531
Qy      140 YserLeuAlaAlaCysAlaPheSerProAsnGlySerPhePheValThrGlySerSerCys 160
Db      532 CTCTTGTGGGGGAGATGCTCTTCTCTCTATGAGAGTTTCTTGTCTAGCTGCTCTCATG 591
Qy      160 sGlyAspLeuThrValTrpAspAspLysMetArgCysLeuHisSerGlyLysAlaHisAs 180
Db      592 TGGGATTTTATCACTGCTGGGATGATTAATAAGAGTGTCTCATATGTAAGAAACACATGA 651
Qy      180 PheGlyIleThrCysGlyAspPheSerSerGlnProValSerAspGlyValGlnGlnGly 200
Db      652 TCTTGATTAATCACTGCTGCAATTTTCTTACAGCCAGTTTCATGATGAGAAACAAGTCT 711
Qy      200 uGlnPhePheArgLeuAlaSerCysGlyGlnAspCysGlnValYsIleTrpIleValSer 220
Db      712 TCACTTTTTCGACTGCAATCATGTGTGTAAGATGTCGAAGTCAAAATTTGATGTTC 771

```

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 1653.r

For more information about this cluster, see <http://www.genoscope.cns.fr/cdna?cs=CS001075D040P1&c=1653.r>.

Location/Qualifiers

(destroyed); Site 2: NCI; RNA source male hippocampus, age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 0.9-4 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 012."

ORIGIN

Alignment Scores:

Pred. No.:	1,07e-106	Length:	1010
Score:	1154.00	Matches:	222
Percent Similarity:	96.968	Conservative:	1
Best Local Similarity:	96.524	Mismatches:	2
Query Match:	56.384	Indels:	5
DB:	4	Gaps:	0

US-10-077-111-13 (1-384) X BM543484 (1-1010)

QY		1	MetValysLeuIleHisThrLeuAlaSerHisGlyAspAspValAsnCysCysAlaPhe	20
Db		186	ATGGGAACCGATTCAACAATTACTGATCATGTGTGACAGTCACTGCCTTC	245
QY		21	SerPheSerLeuLeuAlaThrCysSerLeuAspLysTrpIleArgLeuTySerLeuArg	40
Db		246	TCCCTTTCCCTTGCGTACTTGCTCTCTTGACAAACAAATCGCCTGTACTGGTAGCT	305
QY		41	AspPheThrGluLeuProHisSerProLeuLysPheHisThrTyrAlaValHisCysCys	60
Db		306	GACTTTACTGTAACGCCAACATCTCCATTTCCATGGAATTTCTACTAAGCTGTCCACGCTGC	365
QY		61	CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrThrValLeu	80
Db		366	TGTTTCTCCCTTCACGACATATTTTTGGCATCGTGTTCACAGAAGTGRACACTGCTCTA	425
QY		81	TrpAsnThrGluAsnGlyGlnMetLeuAlaValMetGluGlnProSerGlySerProVal	100
Db		426	TGGAAATCTGAATAATGACACAGATGTGGCAGTGAAGAACAGCTAGTGGCACGCCCTGTG	485
QY		101	ArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyAlaAlaAspGlyThr	120
Db		486	AGGGTTTGGCCAGTTTCCCCAGACTCCAAGTGTGGCATCAGCGCAGCTATGGAAGT	545
QY		121	ValValLeuThrPasnAlaGlnSerTyrLysLeuTyrArgCysGlySerValLysAspGly	140
Db		546	GTGGTTTGTGGAAAGCACAGCTCATACAAATTAATAAGTGTGAAGTTAAAAGATGGC	605
QY		141	SerLeuAlaAlaCysAlaPheSerProAsnGlySerPhePheValThrGlySerSerCys	160
Db		606	TCCTTGGCGGCGATGTGCATTTTCTCTTAATGGAACCTTTGTGTACGTGGCTCTCATGT	665
QY		161	GlyAspLeuThrValTrpAspAspLysMetArgCysLeuHisSerGlyIleVala-HisAs	180
Db		666	GGTATTTTAACAGTGGGAGATGATTAATGAGTGTCTGCATAGGAAAAAGCACACATGA	725
QY		180	PLeuGlyIleThrCysCysAspPheSerSerGlnProValSerAspGlyGluGln-GlyL	200
Db		726	TCTTGGAAATTAACCTGCTGGCGATTTTTTCTTCCAGCGCGATTCTGANTGGAAACAAGGTC	785
QY		200	eu-GlnPhePheArgLeuAlaSerCysGlyGlnAsp-CysGlnValLysIleTrp-IleV	219
Db		786	TTCCAGTTTTTTCCAAATGCGATCATGTGTGAGATTTTGGCAGTCCMAAAATTTGGAAATYG	845
QY		219	aLserPheThrHisIleLeu 225	
Db		846	TTTCTTTTACCCTATCTTA 865	
RESULT 7				
BIF603184			858 bp	mRNA linear EST 07-SEP-2000
LOCUS			603249828F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5301559 5'	
DEFINITION			mRNA sequence.	
ACCSSION			BIF603184	

VERSION	BI603184.1	GI:15496123
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	

REFERENCE
1 (pages 1 to 858)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Butelostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D.

FEATURES

Source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5301559"
/tissue_type="hypothalamus"
/lab_host="Dh10B"
/clone_id="NIH_MGC_96"
/notes="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site 1: BamHI, Site 2: SalI-xhoI
(gtcgag); Oligo-dT primed using primer
5'-TTTTTTTTTTTTVN-3', size-selected for average
insert size 2.3 kb and normalized to R05. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Garnick, in
preparation). Library constructed by M. Brownstein
(NIMH/NHGRI, National Institutes of Health). Note: this is
a NIH_MGC Library."

```

ORIGIN

Alignment Scores:

DB:	4	Gaps:	1
Query Match:	50.988	Indels:	3
Best Local Similarity:	89.964	Mismatches:	13
Percent Similarity:	90.798	Conservative:	2
Score:	1043.50	Matches:	215
Pired, No.:	1.54e-95	Length:	856

US-10-077-111-13 (1-384) X BI603184 (1-858)

QY	Db	QY	Db	QY	Db	QY	Db	QY	Db		
1	MeValIysLeuLeuLeuIleThrLeuAlaSerIleGlyAspSerValIleAsnCyAspAlaPhe	20	151	ATGGTGAAGAAAGATTCAACACTTAGCTGATATAGTGTACAGATGTCAACTGTGTGGCTTC	210	21	SerPheSerLeuLeuAlaThrCysSerLeuAspLysThrIleArgLeuYrSerLeuArg	40	211	TCCCTTTTCCCTCTGGCTACTGTGCTCCCTGGACAAACAAATCCGCTGTACTCGTTCAGT	270
QY	Db	QY	Db	QY	Db	QY	Db	QY	Db		
41	AspPheThrGluLeuProHisSerProLeuLysPheHisThrTyrAlaValHisCysCys	60	271	GACTTACTGAACTGCCACATTTCTCCATTGAAGTTTATTACCTTAGTGGCTGTCCACTGTGCG	330	61	CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrThrValLeu	80	331	TGTTTCTCCCTCCGAGACATATTTGGCAGTCGTTCAACAGATGGTCAACACTGTCTCA	390
QY	Db	QY	Db	QY	Db	QY	Db	QY	Db		
81	TrpAsnThrGluAsnGlyGluMetLeuAlaValMetGluGlnProSerGlySerProVal	100	391	TGGAAATCTGAAATATGACAGATCTGGCAGTGTATGGAACAGCCCTATATGGACGCCCTGTG	450						

QY	101	ArgValCysGlnIlnPheSer-ProAspSerThrCysLeuAlaIAserGlyAlaAlaAspGlyThr	120
Db	451	AGGGTTCCAGTATTTCCCCAGACTCCACGCTTTGGCATCAGGGGCAAGCTGATGAACT	510
QY	121	ValValLeuTrpAsnAlaGlnSerTyrIlyLeuTyrArgCysGlySerValIlyAspGly	140
Db	511	GTGGTTTTGTGGAAAGCAGACAGTCATACAAATATATATAGATGGTAGTGTAAAGATGGC	570
QY	141	SerLeuAlaAlaCysAlaPheSerProAsnGlySerPhePheValThrGlySerSerCys	160
Db	571	TCCTTGGGGGAGTGGCATTTTCTCTTAATGGAACTCTTGTCACTGGCTCCCACTG	630
QY	161	GlyAspLeuThrValTrpAspAspIlyMetArg-CysLeuHisSerGlyIlyAlaHisAs	180
Db	631	GGTAGTTTAAACAGTGTGGATGATACATAGAGGGTGTGCTGATGTGAAAAGCACATGG	690
QY	180	PleuGlyIleThrCysCysAsp---PheSerSerGln-ProValSerAsp-GlyGlnGln	198
Db	691	ATCTGGAAATTACCTGGCTGGCCGATTTTCTTCACAGCAAGTCTCTGATTGGAGACAA	750
QY	199	GlyLeuGln--PhePheArgLeuAlaIAser-CysGly-GlnAspCysGlnValIlyIleTr	217
Db	751	GGCTCTTCAAGTTTTTTCGAACTGGCATCGATGTGGTCCAGATTCGCAAGTCCAAATATG	810
QY	217	PleValSer-PheThrHisIleLeuAlaArgArgThrGlu	230
Db	811	GGATGGGTCTCTTTTACCGATATCTTAGCAAGGGGACACAG	851

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
BI821143	60303503071 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5176086 5', mRNA sequence.	BI821143	BI821143			Homo sapiens (human)	NIH-MGC http://mgc.nci.nih.gov/ , 1 (bases 1 to 792)		National Institutes of Health, Mammalian Gene Collection (MGC)	Unpublished (1999)	Contact: Robert Strausberg, Ph.D. Email: c9gabbs-r@mail.nih.gov
		EST.									Tissue Procurement: Life Technologies, Inc.
											cDNA Library Preparation: Life Technologies, Inc.
											cDNA Library Arrayed By: The I.M.A.G.E. Consortium (LLNL)
											DNA Sequencing by: Incyte Genomics, Inc.
											Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
											Plate: LLM11438 row: 1 column: 07
											High quality sequence ecop: 751.

FEATURES

source

```

1. .792
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5176086"
/lab_host="DH10B"
/clone_1fb="NIH MGC 115"
/name="Organ; site 1: Nott; site 2: ECORV (destroyed); RNA
pcmw-sports; pooled brain, lung, testis; Vector:
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
0.150-0.1 primed and directionally cloned (ECORV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH MGC Library."
```

ORIGIN	
Alignment Scores:	
Pred. No.:	2,27e-94
Score:	1031.50
Percent Similarity:	92.21%
Best Local Similarity:	91.77%
Query Match:	50.39%
Ds:	4
US-10-077-111-13 (1-384) x B1821143 (1-792)	
Length:	792
Matches:	212
Conservative:	1
Mismatches:	10
Indels:	9
Gaps:	1

Oy	1	MetValylLeuLeuLeuIshThrLeuAlaAspIscglYaspAPValAenCYsAlaPhe	20
Db	118	AtGATAAACTGATTACACATTACTATATAGTGACATGTCAACTGCTGCTTC	177
Oy	21	SerPheSerLeuLeuAlaThrCYsSerLeuAspYsthrIleArgLeuTYrSerLeuArg	40
Db	178	TCTTTTTCCTCTGGGTACTGCTC---CTGACAAACAAATTCGCGCTGTAAGT	234
Oy	41	AspPheThrGluLeuProHisSerProLeuLYsPheIsthrTYrAlaValHisCYsGys	60
Db	235	GACATTACGAACCTGCACATCTTCATTGAAGTTCAACTCTATGCTGCACACTGTC	294
Oy	61	CysPheSerProSerGlyHisIstLeuAlaSerCYsSerThrAspGlyThrTrpValLeu	80
Db	295	TGTTTCTCCCTTCAGACATAT-TTGGCATCTGGTTCACAGATGGTACACATGCTCA	353

OY	8	TRPAsnThGluAsnGlyValMetCysLeuAlaValMetGluGlnProSerGlySerProVal	100
Db	354	TGGAAATACAGAAATGACACAGATGCTGGCAGTGAAGAACGCTTAAGTGGCAAGCCCTGTG	4133
OY	101	ArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyValAlaAspGlyThr	1200
Db	414	AGGGTTGGCAAGTTTCCCGACAGATCCACAGTGTTGGCATCAGGGGACAGCTGATGAAACT	4733
OY	121	ValValLeuThrAsnAlaGlnSerPyrtyrLeuPyrtyrArgCysGlySerValIlyAspGly	1400
Db	474	GTCGTTTTGGAAATGACACAGTCACAAATTAATAGATGTGTAGTGTTAAAGATGGC	5333
OY	141	SerLeuAlaAlaCysAlaPheSerProAsnGlySerPhePheValThrGlySerSer-Cy	1600
Db	534	TCCTTGGCGGCAATGTCATTTTCTCTAATAGAAAGCTTTTGTGCATGGCTCCTCAATG	5933
OY	160	ArgIAspLeuThrValTTPAspAspLysMetArgCysLeuHisSerGlu-LysAlaHis	1800
Db	594	TGGTGATTTAAACAGTGTGGGATGAATAAATAGGGTGTGCATGTGTAAATAAAGCTCATG	6533
OY	180	AspLeuGlyIleThr-CysCysAspPheSerSerGlnProValSer-AspGlyGluGlnG	1999
Db	654	ATCTTGGAAATTACCTGCTGTCGCAATTTTCTTCACAGCCAGATTTTGTATGGAAACACAGG	7133
OY	199	LysGlnPhePheArg-LeuAlaSerCys-GlyGlnAspCysGlnValLys-IleTyrP	2188
Db	714	TCTTCAGTTTTCGCACTGGCATCATGTGGGTCCAGATGGGATGCATGCACAAATTTGGAC	7733
OY	218	eValSerPheThrHisIle 224	
Db	774	TGTTACTTTTACCCATATT 792	

RESULT 9

LOCUS DEFINITION ACCESSION VERSION KEYWORDS SOURCE ORGANISM	CN793033 CN793033 41278990 BARC 8BOV Bos taurus CDNA clone 8BOV_46H02 5' , mRNA sequence. CN793033 CN793033.1 GI:47689013 EST.	728 bp mRNA linear EST 26-MAY-2000
--	--	---

Bos taurus (cow)
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eumleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovinae; Bos.

Qy	234	lyscglnpethrctgluaspstpsrcstlucgluvalvalserthrtrpplencg	251
Db	654	---AAAGACACTTAAGTGGGACATGGCCCTTCTCT--TGACCTTGCT	701
RESULT 12			
LOCUS	CK306752		
DEFINITION	SB02035A1D03.f1 normalized Keck-Tagu Library SB02 Taeniopygia		
ACCESSION	CK306752		
VERSION	CK306752.1	GI:44816326	
KEYWORDS	EST.		
SOURCE	Taeniopygia guttata		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Archosauiria; Aves; Neognathae; Passeriformes; Estrildidae;		
	Estrildinae; Taeniopygia.		
	1 (bases 1 to 839)		
REFERENCE	Clayton,D.F., Arnold,A.P., Ball,G.F., Brenowitz,E., George,J.M.,		
AUTHORS	Mello,C.V., Wade,J., Replogle,K., Lewin,H., Band,M., Hernandez,A.		
	and Liu,L.		
	The Songbird Neurogenomics Initiative: An Evolving Public Resource		
	for Study of Genes, Brain, and Behavior		
	Unpublished (2004)		
JOURNAL	Contact: David P. Clayton		
COMMENT	University of Illinois		
	B107 CLSL, 601 S. Goodwin, Urbana, IL 61801, USA		
	Tel: 217 244 3668		
	Fax: 217 244 1648		
	Email: dclayton@uiuc.edu		
	Base Calling/Quality Scores: PHRED from Washington University		
	Genome Center.		
	Vector Trimming: Cross match from Washington University Genome		
	Center PHRAP suite. Low quality bases (phred score < 20) were		
	trimmed from both ends of the sequence by an in-house script.		
	This sequence is vector free and at least 200 bp in length. Funded		
	by PHS grant # R01 NS045264, 'Songbird Neurogenomics Initiative.'		
	PCR Primers		
	FORWARD: TAATACGACTCACTATAGG(T7)		
	BACKWARD: ATTTACCTCCTAATAG(T3)		
	Insert Length: 839 Std Error: 0.00		
	Plate: SB02035A1 row: D column: 03		
	Seq primer: TAATACGACTCACTATAGG (T7)		
	High quality sequence stop: 839.		
FEATURES			
source	1..839		
	location/Qualifiers		
	/organism="Taeniopygia guttata"		
	/mol_type="mRNA"		
	/db_xref="taxon:59729"		
	/clone="SB02035A1D03.f1"		
	/tissue_type="brain"		
	/dev_stage="late embryo, post-hatch days 1, 10, 20, 45,		
	and adult (pooled)"		
	/lab_host="DH10B"		
	/clone_lib="normalized Keck-Tagu Library SB02"		
	/note="Organ: brain; Vector: pBS II SK(+); Site 1:		
	EcoRI(5' side of insert); Site 2: NotI (3' side of		
	insert). The library was constructed and normalized as		
	described by Bonaldo, M.F., Lennon, G. and Soares, M.B.		
	(1996), Genome Research 6(9): 791-806. An identifying tag		
	was added at the 3' during cDNA synthesis:		
	insertAAAAAAAAAAAAAAAAAATGCA."		
ORIGIN			
Alignment Scores:			
Pred. No.:	1.24e-81	Length:	839
Score:	907.00	Matches:	161
Percent Similarity:	84.00%	Conservative:	28
Best Local Similarity:	71.56%	Mismatches:	36
Query Match:	44.31%	Indels:	0
DB:	7	Gaps:	0
US-10-077-111-13(1-384) x CK306752 (1-839)			

OY		MetValylSerLeuIleHisThrLeuAlaAspArgGlyaspPheValAsnCySAlaPhe	20
Dd		ATGGCGACATTAAATTCACACTTTAGCAGATATAAATGAATGTTAATTAATCACTGCCTTCC	149
OY		SerPheSerLeuLeuAlaIathrCySeSerLeuAspLyThrIleaArgLeuTrSerLeuArg	40
Dd		TCAACTTGTCGCCCTGGCTACATGTTCTCTGGACAACAAATTCGCCGTATATTCCTTTAGACC	209
OY		AspPheThrGluLeuProHiseSerProLeuLysPheHisThrTyrlValAlaHiAcSecys	60
Dd		AAC TTC GGT CAG CTC CCC GT ACT CC CG T AC GG AG GC TA CA CG G TC A CG CC TG CAT CTG C GC	269
OY		CysPheSerProSerArgIyHisIIeLeuAlaSerCysSerThrAspGlyThrrTheValneu	80
Dd		TGCTTCTCGCGGTGGAGGCGCTGTGTGTGCTCTGTGCCTCCAGCATTCGACCTCGGTGCTG	329
OY		TPasanthrgluaanglglnMetLeuAlaValMetgluginProsergilySerProval	100
Dd		TGGGACACCGCGATGGCCGACAGGCTGGCCGTGTGGAGCACGCCGGGGCCAGCCCTGTC	389
OY		ArgValCysglinPheSerProAspSerThrCysleuAlaserglyalialaspglythr	120
Dd		AGGCGTCCGCCCTCTCGCCGAGAAGCCGCCCTATCTCTTGGCAGGGGCGACGCGATGGCAGC	449
OY		ValValleuThrPaanaIagInseTrytyrylleutyrargcysgliserValiyasaSpoly	140
Dd		GTTGTTCTCTGGAACGCGCATCTCAATGAATCTTAACGATTTGGGAAAATTMAAGCTGT	509
OY		SerLeuAlalaCySaIaPheSerProasnlglycerPhenavalThnglysSerSerCys	160
Dd		TCTTTGAATGGCTTGTCATTTTTCTCCCAGTAGAAACTCTTTGTACATGATCATCAAGT	569
OY		GlyaspLeuthrValITrpaspApIymEckarGcySleuHisSerGlubysAlaHiasep	180
Dd		GGTGAATTMAACATTGGGATGATAAATGAATGAAGCCGTMTAATGAAAAAGCACATGAT	629
OY		LeuglyIleThrCysCaSPASPheSerSerGlnProvalSerAspGlyGIUIngIGlyeu	200
Dd		CTTGCGCTTACTCTGCTGTGATATTTCTTTCACATCAAGTATTCGATAGGAAAATGAGATGC	689
OY		GlnphePhearGleualaJerCySGlyGlnaspCysglinaVallysileTriplieValaser	220
Dd		AAATAC TTCCAG ATGGCTTCTCTGTGACCAAAGATATCATCAACTGTGCTTATTTTG	749
OY		PheThrHisIleLeu	225
Dd		TTTGACAGATTTCTTA	764
RESULT_13			
LOCUS	CKJ92250	808 bp mRNA linear EST-25-FEB-2004	
DEFINITION	AGNCOURT18674252 NIH-MGC-230 Mus musculus cDNA clone IMAGE:30845589.5'		
ACCESSION	CKJ92250		
VERSION	CKJ92250.1 GI:42804246		
SOURCE	Estr.		
ORGANISM	Mus musculus (house mouse)		
JOURNAL COMMENT	Mus musculus Buxarayota; Metazoa; Chordata; Craniata; Vertebrata; Eutelestomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 808) NIH-MGC http://mgc.nci.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Daniela S. Gerhard, Ph.D. Office of Cancer Genomics National Cancer Institute / NCI Bldg. 31 Rml0A07 Bethesda, MD 20892 Email: gcgaops@mail.nih.gov Tissue Procurement: Shioho Kimura/Abeuchi Yamada, (NCI,CCR) cDNA Library Preparation: Express Genes GenBank Accession: U01616 CDNA library Arrayed by: The I.M.A.G.E Consortium (LMNL)		

Db 2 GCTCTGAGAGACATATTTGGCATGCTGTTCAACAGATGGTACCAGTCTTATGAAAT 61
Qy 83 ThrGluanglYgImeLleuAlaValMetGluGlnProSerGlySerProValArgVal 102
Db 62 ACTGAATAATGACAGATGCTGGCAGTATGAGAACAGCTTACTGACGCCCTGTAGAGGTT 121
Qy 103 CysGlnPheSerProAspSerThrCysLeuAlaSerGlyAlaAlaAspGlyThrValVal 122
Db 122 TGCAGATTTCCTCCAGACTCCAGCTGTTGGCATCAGGGGAGCTGATGAGAACTGTGGTT 181
Qy 123 LeuThrPsnAlaGlnSerThrValLeuThrArgCysGlySerValValAspGlySerLeu 142
Db 182 TTGGGATGACAGATGATTAATAATGATATGATGATGATGATTAAGATGATGATGATGAT 241
Qy 143 AlaAlaCysAlaPheSerProAspGlySerPhePheValThrGlySerSerCysGlyAsp 162
Db 242 GCGGCAATGTCATTTTCTCTTAATGAGAGCTTTCTTGTACCTGGCTCTCATGTGGTAT 301
Qy 163 LeuThrValITrPaAspAspLysMetArgCysLeuHisSerGlyLysAlaHisAspLeuGly 182
Db 302 TTACAGATGTCGATGATTAATAATGATGATGATGATGATGATGATGATGATGATGAT 361
Qy 183 IleThrCysCysAspPheSerSerGlnProValSerAspGlyGluGlnGlyLeuGlnPhe 202
Db 362 ATTACCTGCTGCGATTTTCTTCAACAGCAGCTTCTGATGAGAAACAAGTCTTCAGATT 421
Qy 203 PheArgLeuAlaSerCysGlyGlnAspCysGlnValLysIleTrrPileValSerPheThr 222
Db 422 TTTCGACTGGCATCATGTGTCAGAGATTTGCCAAGTCAAAATTTGGATGTTCTTTTACC 481
Qy 223 HisIleLeu 225
Db 482 CATATCTTA 490

RESULT 15
BIT59505 696 bp mRNA linear EST 25-SEP-2001
LOCUS 603046888F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5187309 5',
DEFINITION mRNA sequence.
ACCESSION BIT59505
VERSION BIT59505.1 GI:15751083
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 696)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rgabbe-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLU)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLU at:
http://image.llnl.gov
Plate: L14M11467 row: 0 column: 22
High quality sequence stop: 696.
Location/Qualifiers
1..696
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5187309"
/lab_host="MDH10B"
/clone_lib="NIH_MGC_116"
/note="Organ: pooled colon, kidney, stomach; Vector:
pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of 3 colons, age 26 yo male, 49 yo
female, 71 yo male colon; 46 yo male kidney, and pool of 2

stomachs, 62 yo male and 70 yo female. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.4 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
023. Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.:	5,26e-79	Length:	696
Score:	880.00	Matches:	161
Percent Similarity:	100.00%	Conservative:	2
Best local Similarity:	98.77%	Mismatches:	0
Query Match:	42.99%	Indels:	0
DB:	4	Gaps:	0

US-10-077-111-13 (1-384) x BIT59505 (1-696)

Qy 63 SerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrThrValLeuTrpAsn 82
Db 2 GCTCTGAGAGACATATTTGGCATGCTGTTTCAACAGATGGTACCAGTCTTATGAAAT 61
Qy 83 ThrGluanglYgImeLleuAlaValMetGluGlnProSerGlySerProValArgVal 102
Db 62 ACTGAATAATGACAGATGCTGGCAGTATGAGAACAGCTTACTGACGCCCTGTAGAGGTT 121
Qy 103 CysGlnPheSerProAspSerThrCysLeuAlaSerGlyAlaAlaAspGlyThrValVal 122
Db 122 TGCAGATTTCCTCCAGACTCCAGCTGTTGGCATCAGGGGAGCTGATGAGAACTGTGGTT 181
Qy 123 LeuThrPsnAlaGlnSerThrValLeuThrArgCysGlySerValValAspGlySerLeu 142
Db 182 TTGGGATGACAGATGATTAATAATGATATGATGATGATGATTAAGATGATGATGATGAT 241
Qy 143 AlaAlaCysAlaPheSerProAspGlySerPhePheValThrGlySerSerCysGlyAsp 162
Db 242 GCGGCAATGTCATTTTCTCTTAATGAGAGCTTTCTTGTACCTGGCTCTCATGTGGTAT 301
Qy 163 LeuThrValITrPaAspAspLysMetArgCysLeuHisSerGlyLysAlaHisAspLeuGly 182
Db 302 TTACAGATGTCGATGATTAATAATGATGATGATGATGATGATGATGATGATGATGAT 421
Qy 183 IleThrCysCysAspPheSerSerGlnProValSerAspGlyGluGlnGlyLeuGlnPhe 202
Db 362 ATTACCTGCTGCGATTTTCTTCAACAGCAGCTTCTGATGAGAAACAAGTCTTCAGATT 421
Qy 203 PheArgLeuAlaSerCysGlyGlnAspCysGlnValLysIleTrrPileValSerPheThr 222
Db 422 TTTCGACTGGCATCATGTGTCAGAGATTTGCCAAGTCAAAATTTGGATGTTCTTTTACC 481
Qy 223 HisIleLeu 225
Db 482 CATATCTTA 490

Search completed: February 5, 2005, 14:52:37
Job time : 4214 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 5, 2005, 05:07:50 ; Search time 246 Seconds
(without alignments)
8460.748 Million cell updates/sec

Title: US-10-077-111-12

Perfect score: 1272

Sequence: 1 gaatcgcttcacacgctgcg.....acacaaagtaagaattc 1272

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/5A COMB.seq:*

2: /cgn2_6/ptodata/1/ina/5B COMB.seq:*

3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*

5: /cgn2_6/ptodata/1/ina/PTUS COMB.seq:*

6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	964.2	75.8	1844	4	US-09-620-312D-768
2	49.6	3.9	7218	1	US-08-232-463-14
3	45.8	3.6	7218	1	US-08-232-463-14
4	44.8	3.5	1221	3	US-08-965-600-2
5	44.8	3.5	1221	3	US-09-489-506-2
6	43.8	3.4	1141	4	US-09-806-708B-22
7	43.6	3.4	1141	4	US-09-806-708B-22
8	43.6	3.4	937	4	US-10-101-464A-251
9	40.4	3.2	162450	1	US-08-188-582-17
10	40	3.1	2152	1	US-08-646-715-17
11	40	3.1	246240	2	US-08-724-394A-20
12	40	3.1	246240	2	US-08-724-394A-20
13	40	3.1	246240	2	US-08-724-394A-21
14	39.6	3.1	5152	3	US-09-690-364-10
15	39.6	3.1	1664976	4	US-08-916-421B-1
16	39.6	3.1	1664976	4	US-09-692-570-1
17	39.6	3.0	2085	2	US-08-283-917-8
18	38.2	3.0	2085	2	US-08-961-716-8
19	38.2	3.0	2142	4	US-09-614-221A-331
20	37.8	3.0	601	4	US-09-949-016-138114
21	37.8	3.0	601	4	US-09-949-016-138114
22	37.8	3.0	601	4	US-09-949-016-138114
23	37.8	3.0	601	4	US-09-949-016-138114
24	37.8	3.0	112705	4	US-09-949-016-138116
25	37.6	3.0	1464	4	US-09-248-796A-1456
26	37.4	2.9	29357	4	US-09-949-016-16676
27	37	2.9	732	4	US-09-248-796A-2341

28	36.8	2.9	1611	3	US-09-302-769-13	Sequence 13, Appl
29	36.6	2.9	3127	4	US-09-710-279-3502	Sequence 3502, Ap
30	36.6	2.9	6968	4	US-09-710-279-759	Sequence 759, Ap
31	36.6	2.9	7215	3	US-09-134-001C-627	Sequence 627, Ap
32	36.4	2.9	3081	4	US-09-949-016-2460	Sequence 2460, Ap
33	36.2	2.8	1566	4	US-09-248-796A-9260	Sequence 9260, Ap
34	35.8	2.8	389	4	US-09-270-767-26536	Sequence 26536, A
35	35.8	2.8	832	4	US-09-621-976-2813	Sequence 2813, Ap
36	35.8	2.8	899	4	US-09-370-767-11026	Sequence 11026, A
37	35.8	2.8	1743	4	US-09-248-796A-11015	Sequence 11015, A
38	35.6	2.8	1182	4	US-09-248-796A-4568	Sequence 4568, Ap
39	35.6	2.8	4892	4	US-09-976-594-167	Sequence 167, Ap
40	35.6	2.8	5158	4	US-09-023-655-1347	Sequence 1347, Ap
41	35.6	2.8	640681	4	US-09-790-988-1	Sequence 1, Appl
42	35.6	2.8	640681	4	US-09-790-988-1	Sequence 1, Appl
43	35.4	2.8	2417	4	US-09-254-776B-3	Sequence 3, Appl
44	35.4	2.8	3747	3	US-09-690-364-17	Sequence 17, Appl
45	35.4	2.8	7028	4	US-09-949-016-4194	Sequence 4194, Ap

ALIGNMENTS

RESULT 1

US-09-620-312D-768

Sequence 768, Application US/09620312D

Patent No. 6569662

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom

APPLICANT: Liu, Chenghua

APPLICANT: Asundi, Vinod

APPLICANT: Zhang, Jie

APPLICANT: Ren, Feiyang

APPLICANT: Chen, Rui-hong

APPLICANT: Zhao, Qing A.

APPLICANT: Wehrman, Tom

APPLICANT: Xue, Aidong J.

APPLICANT: Wang, Yonghong

APPLICANT: Wang, Jian-Rui

APPLICANT: Zhou, Ping

APPLICANT: Ma, Yungqing

APPLICANT: Wang, Dunru

APPLICANT: Wang, Zhiwei

APPLICANT: John Tillinghast

APPLICANT: Dmanac, Radoje T.

TITLE OF INVENTION: No. 6569662el Nucleic Acids and

FILE REFERENCE: 784CIP28

CURRENT APPLICATION NUMBER: US/09/620,312D

PRIOR FILING DATE: 2000-07-19

PRIOR APPLICATION NUMBER: 09/552,317

PRIOR FILING DATE: 2000-04-25

PRIOR APPLICATION NUMBER: 09/488,725

NUMBER OF SEQ ID NOS: 1105

SOFTWARE: pc_FL gene Version 1.0

SEQ ID NO 768

LENGTH: 1844

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (164) .. (1594)

US-09-620-312D-768

Query Match 75.8%; Score 964.2; DB 4; Length 1844;

Best Local Similarity 81.8%; Pred. No. 1.9e-307;

Matches 1252; Conservative 0; Mismatches 3; Indels 276; Gaps 1;

QY 12 TCACCTGCGCGGACCTGACCCGCGCGGACCTTGAGCGGATCCGCGCG 71

Db 65 TCACCTGCGCGGACCTGACCCGCGCGGACCTTGAGCGGATCCGCGCG 124


```

; IMMEDIATE SOURCE:
; CLONE: PTZgpc-Flis
US-08-232-463-14

Query Match      3.9%; Score 49.6; DB 1; Length 7218;
Best Local Similarity 8.3%; Pred. No. 0.0001;
Matches 34; Conservative 200; Mismatches 174; Indels 0; Gaps 0;

QY      4 TTGCGCTTCACTGCAGCAGTGACCCGCCCGGTGGGCACTTGAAAGCGGAT 63
       :: : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      1102 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1161
QY      64 CCCC GCCGCCCCCGCTCCGTCGACGCGTTTCTTCAATAAAGAACAATGGGAACCTGA 123
       :: : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      1162 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1221
QY      124 TTCACACATTAGCTGATCATGTGACGATGTCACACTGCTGTGCTTCCTTCCCTCT 183
       :: : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      1222 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1281
QY      184 TGCTACTTGTCTCTTGGA CA A A C A A T T C G C T G T A C T T A C T T A C T G A C 243
       : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      1282 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1341
QY      244 TGCCACATTCOCATGAAGTTTCACTTAACCTAGCTGCTGCTGCTTCGCCCTT 303
       : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      1342 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1401
QY      304 CAGCACATATTTGGCATGCTGTTCACAGATGTACCACACTGCTCTATGAACTGA 363
       : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      1402 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYGACCAAAATCTTATCTCTTTAACT 1461
QY      364 ATGACAGATGCTGCACTGATGAACAGCCTTAGTGAGCCCTTGGA 411
       ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      1462 ACTTSCATAGATAGGTAAATTACAGATGATGCTCAATGCGCTTTTGA 1509

RESULT 3
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHELFINGER, F.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232.463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935.313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:

```

```

TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TRLEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZSpC-Fls
US-08-232-463-14

Query Match      3.6%, Score 45.8; DB 1; Length 7218;
Best Local Similarity 9.8%, Pred.No. 0.0018;
Matches 47; Conservative 215; Mismatches 217; Indels 0; Gaps 0;

686 AGTTTCGATGAGAGACAAGGCTTCGATGTTTTTCGACTGGCATCATGTGGTCAGATTG 745
1539 AGCATCGCTTCAGACCGCATCTATTCGATTTCAAAAACGGCATGTAGGCTACTGTAA 1480
746 CCAAGTCAAAATTGGATTGTTCTTTTACCATATCTTAGACAGCGCACAGAACTCA 805
1479 TTACCTATCTATGCAAGTAGTAAAGATVGAAGAATTGGTACRRRRRRRRRRRRRR 1420
806 GCTAAGCAATTACCGAAGATTGGTCAGAGAGAGTGTCTCAACATGGCTTTTGACA 865
1419 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1360
866 AGATTTAAAGATTTGTTGGTATTTTCAAGATGAATTAACATGATGAAAGAACTGTT 925
1359 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1300
926 GAATCTTCAAAAGAAAGTGTGGCTGATGATTTGAAATTTGAATCTAGACTGCGTAG 985
1299 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1240
986 TAAAGTGTGAGAAAATTGAGAGCTCAGACCAAGTTAAATCCCTTTCTTAGCAAT 1045
1239 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1180
1046 TCCGTAGAAATTTATGTGCCAATAACTAGAGAACTTATGAGAAATCGGTCATGCATC 1105
1179 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1120
1106 AGATGCGTATTCATATGAAAGAGCAATGAAATTTGATCGCAAAAAGAAACGTA 1164
1119 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRATCGCA 1061

RESULT 4
US-08-965-600-2
Sequence 2, Application US/08965600
Patent No. 6077688
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Corley, Neil C.
APPLICANT: Shah, Purvi
TITLE OF INVENTION: NEW TRANSDUCIN BETA-1 SUBUNIT
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PaeISeq for windows Version 2.0

```

```
/
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/965,600
/ FILING DATE: Herewith
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Billings, Lucy J.
/ REGISTRATION NUMBER: 36,749
/ REFERENCE/DOCKET NUMBER: PF-0416 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 650-855-0555
/ TELEFAX: 650-845-4166
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1221 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: KIDNOT02
/ CLONE: 194046
/
US-08-965-600-2
```

Query Match 3.5%; Score 44.8; DB 3; Length 1221;

Best Local Similarity 49.3%; Pred. No. 0.0011; Matches 146; Conservative 0; Mismatches 147; Indels 3; Gaps 1;

```
QY 328 CAACAGATGATGACCTGCTCTATGGAATACTGAAAATGACAGATGCTGGCAGTATG 387
DB 579 CCATAGATGGAATCATCATATTTTGTATGCACTGGAATACTTCGATACCTGG 638
QY 388 AACAGCTTAGTGACGCTCTGAGGTTGGCCAGTTTCCCAAGCTCCAGCTGTTGG 447
DB 639 AAGGCC---ATGCCATGCCCATTCGCTCTTGAACCTTTCCCGGACTCCAGCTCCTTG 695
QY 448 CATCAGGGGACGCTAGTGAAGTGTGTTTGTGATGACAGTCAATCAATTAATA 507
DB 696 TCACCTCTTCAGATGATGCTTACATCAAGTCTATGATGTAACAATGCCAATTTGGCTG 755
QY 508 GATGTGATGATGTTAAAGATGCTCCTTGGCGGACATGTCATTTCTCTAATGAAAGCT 567
DB 756 GCACGCTGAGCGGCATGCTCCTGGGTGCTGAACGTGATTCGTGCGATGACATC 815
QY 568 TCTTTGTCACTGCTCTCTCATGTGTGATTTAAACAGTGTGGGATTAATAATGAG 623
DB 816 ACTTTGTTCCAGTTGCTGTGACMAAAGTGTAAAGTTGGGATGTTGGAACGAG 871
```

RESULT 5

```
US-09-489-506-2
/ Sequence 2, Application US/09489506
/ Patent No. 6465619
/ GENERAL INFORMATION:
/ APPLICANT: Bandman, Olga
/ APPLICANT: Lal, Preeti
/ APPLICANT: Corley, Neil C.
/ TITLE OF INVENTION: NEW TRANSDUCIN BETA-1 SUBUNIT
/ NUMBER OF SEQUENCES: 3
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Incyte Pharmaceuticals, Inc.
/ STREET: 3174 Porter Drive
/ CITY: Palo Alto
/ STATE: CA
/ COUNTRY: USA
/ ZIP: 94304
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
```

```
/
/ SOFTWARE: FastSeq for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/489,506
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/965,600
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Billings, Lucy J.
/ REGISTRATION NUMBER: 36,749
/ REFERENCE/DOCKET NUMBER: PF-0416 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 650-855-0555
/ TELEFAX: 650-845-4166
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1221 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: KIDNOT02
/ CLONE: 194046
/
US-09-489-506-2
```

Query Match 3.5%; Score 44.8; DB 3; Length 1221;

Best Local Similarity 49.3%; Pred. No. 0.0011; Matches 146; Conservative 0; Mismatches 147; Indels 3; Gaps 1;

```
QY 328 CAACAGATGATGACCTGCTCTATGGAATACTGAAAATGACAGATGCTGGCAGTATG 387
DB 579 CCATAGATGGAATCATCATATTTTGTATGCACTGGAATACTTCGATACCTGG 638
QY 388 AACAGCTTAGTGACGCTCTGAGGTTGGCCAGTTTCCCAAGCTCCAGCTGTTGG 447
DB 639 AAGGCC---ATGCCATGCCCATTCGCTCTTGAACCTTTCCCGGACTCCAGCTCCTTG 695
QY 448 CATCAGGGGACGCTAGTGAAGTGTGTTTGTGATGACAGTCAATCAATTAATA 507
DB 696 TCACCTCTTCAGATGATGCTTACATCAAGTCTATGATGTAACAATGCCAATTTGGCTG 755
QY 508 GATGTGATGATGTTAAAGATGCTCCTTGGCGGACATGTCATTTCTCTAATGAAAGCT 567
DB 756 GCACGCTGAGCGGCATGCTCCTGGGTGCTGAACGTGATTCGTGCGATGACATC 815
QY 568 TCTTTGTCACTGCTCTCTCATGTGTGATTTAAACAGTGTGGGATTAATAATGAG 623
DB 816 ACTTTGTTCCAGTTGCTGTGACMAAAGTGTAAAGTTGGGATGTTGGAACGAG 871
```

RESULT 6

```
US-09-806-708B-22
/ Sequence 22, Application US/09806708B
/ Patent No. 6784342
/ GENERAL INFORMATION:
/ APPLICANT: The University of British Columbia
/ TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
/ FILE REFERENCE: 4810-58741
/ CURRENT APPLICATION NUMBER: US/09/806,708B
/ CURRENT FILING DATE: 2001-04-03
/ PRIOR APPLICATION NUMBER: US 60/147,133
/ PRIOR FILING DATE: 1998-08-04
/ NUMBER OF SEQ ID NOS: 23
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 22
/ LENGTH: 1141
/ TYPE: DNA
/ ORGANISM: Artificial sequence
/ FEATURE:
/ NAME/KEY: promoter
/ LOCATION: (1)..(1141)
```


OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. PAB1 promoters
US-09-806-708B-22

Query Match 3.4%; Score 43.8; DB 4; Length 1141;
Best Local Similarity 9.6%; Pred. No. 0.0023;
Matches 67; Conservative 279; Mismatches 341; Indels 9; Gaps 1.

QY	CGGATCGGCAATTTTCTCTAATGGAACCTTTTGTCAGTGGCTCCCATGGTGATTT	538
Db	53 MSKRKMTWAMMYCKYRRYYNNKSHWKGWYKKMYVCAANNSTSBYYHARRMKDXTAYBM	112
QY	TAACAGTGTGGATGATTAATAATGAAGTGTCTGCATAGTGAATAAGACATGATCTTGAA	598
Db	113 TMTKMGCTGRRHHYWRBAMBBDYDHHYYTANNNATTTTCCMDDDKRTTMMWKNNNA	172
QY	TTACCTGTCTGGATTTTTCCTCACAGCAATTTTGATGAGAAACAAGTCTTCAGTTT	658
Db	173 TGMDDDTKYHHMNNNGCBTVMYWRKYTRDWSBKRMYGMBWMKMGSDYTYWYVWMD	232
QY	TTGCACATGGCATCATGTGTGAGAGATTGCCAAGTCAAAATTTGGATTGTCTTTACC	718
Db	233 MCKKKYVRWVLT-----RGRRRNYVAMBTAHRRRYNNGWTBAMAYRRMTNNNNN	283
QY	ATATCTTAGCAGGGGACAGACAATCAGCTGAAGCAATTTAACCAATATGGTCAGAG	778
Db	284 NAKAMCKRAKYGWNRBAVNSTCTTWSKTTKVRFSCAANNCRAGDAKKHMKWSAAM	343
QY	AGTCGTCTCAACATGCGTTGTGTGCACAGATTTAAAGAATCTTGTTGGTATTTTCAAG	838
Db	344 GVVNNNNNNNNWYTKAARHBARWDMVHSAKKMHANAHAHSRKMTBPKYKRTWNNNG	403
QY	TGAATTAACATTGATGAGAAAAGAATGTTGAATCTTACAAAAGAAAGTGTGCTGATTT	898
Db	404 TTMKKRWAMWYMKDMDBGTYYNNNNNGRTYYGWTCKKMKWYTYKWKANNCKPRAMDHK	463
QY	TGAAATTTGATCTCTAGSAGTCTGCTATTAAGTCTGAGAGAAAATTAAGACTCAGAG	958
Db	464 TCTNNNTTMMCKYUWNNCYWKSMTNGKSHBAAAVYWMWMMRYAHANNNNMDWYWK	523
QY	CSAAGGTAAATCCCTTTCTCAGAGAAATCTCGATTAATATGCTCAATTAATAGAG	1018
Db	524 ACTYKKBVCGKMMNNYAAWTTKSSKNYTSRYRKTKNNSPRKBSDTSMGRANNYAAAB	583
QY	AACCTATGAAGAAGTCCGCTCATCGCATCAGATGGCTATTCATATGAAAAGAACCAATG	1078
Db	584 HYGYKMNTRMBWSHTHBEBRAGAHHYMMBMYBAKCHCMKAWYAKKYAGAGGSGNNNN	643
QY	AAAAATGGATCAGCAAAAAGAAAGCTACAAGTCCCATGACAAATTTGTCTTCTTAG	1138
Db	644 NNNNNNNNNNNATCARDDYAAASRYAAMAAKMYUYKEAANNAAYYTHANNWGCNNATD	703
QY	CGGTACTTACACCAATAGACTCGAATAATGGCCA	1233
Db	704 TTRTMMKNNNNNAGTWKNNNNNNNAKNSAAANVAA	739

```

RESULT 7
US-09-806-708B-22/c
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-585741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA

```

```

: ORGANISM: Artificial sequence
:
: FEATURE:
: NAME/KEY: promoter
: LOCATION: (1)..(1141)
:
: OTHER INFORMATION: consensus sequence of A.c., L.a., and B.n. FAEI promoters
US-09-806-708B-22

```

Query Match	3.4%;	Score 43.6;	DB 4;	length 1141;
Best Local Similarity	13.2%;	Pred. No. 0.0027;		
Matches	77;	Conservative 192;	Mismatches 309;	Indels 5; Gaps 1

QY	690	TCGATGAGAAACAAGCTCTTCAGTTCTTTTTCGATGCGATCATGTGCTAGATTGCCAA	749
Db	1100	TKSATSATGAWTTHAKGATMCCTWTYTGTNRRCRCATTAMRTYTTESNAMSATGEBM	10411
QY	750	GTCMAAATTGGATTGTTCTTTTCCATATCTTAGCAAGCGCAGCAATCAGCTG	809
Db	1040	WTMKWTATKXRTIAYAWCAMRNNNNMCATNGVAKSCATNNAMMYATTBMAVAAAKWAR	981
QY	810	AAGCAATTACCGAAGATTGGTCAGAGAGCTCGTCTCAACATGCTTTGTGCACAAGT	869
Db	980	WAGNNMMYGAAGAGNKGCMAAAMTWBMAWDTAGKCCNNNNNTTWDVRMAKAKXNN	921
QY	870	TTAAGAATGCTGTGGTATTTTCAAGATGATTAACATGATGGAGAAAGAACTGTGAAT	929
Db	920	NNNAVTACTNRAATTNNKMAITHMKWTHGAHSKRTHHTCRKTYNNNNNARYTW	861
QY	930	CTTACAAAGAAAGTCTGGCTGATGATTTGAAAAATTGAATCTCTAGACCTGCGTAGTAA	989
Db	860	YHHAARRMNAWMTRTNNNNNNNNNNACBMTRTWAMBKHSKCNNNNNNNNNNNTWC	801
QY	990	GTCGTGAGAAAATTGGAAGCTCAGACCAAGGTTAAATCCCTTTCTTCAGAATTCT	1049
Db	800	HYTTANABCYRANNNNAARMAATCANNYHAAVTTHTBWCYKTNWATWMDMTTMBT	741
QY	1050	GATGAATTTAATGATGCGAATGACAGAGAACTTAGAAAGATCGGTCATGCGATCGAT	1109
Db	740	TTTNNMTTSTNNNTNNNNNNMMATCTNNNNNNMMKAYAHATNNCGMW-----NNTDART	686
QY	1110	GGCTATTTCATATGAAAAAGAAAGCAATGAAAAATTGATCAGCAAAAAGAAACGTACAAGT	1169
Db	685	NNTTVMRRMWTNTKRTWYSTTTRHHYTGATNNNNNNNNNNNNNNNSCCTBTBMTBWB	626
QY	1170	CCCATGACAAATCTTGTCTTCTTCCTTCAGCGGTACTTACCAAAATAGACCTCTGAAATG	1229
Db	625	TMKDEGTVTKKVKWBDTCTTYUDVADSWVWYANMBRCRDVYTYTNNTYCKSAHSYW	566
QY	1230	GCCATCAATAGATGGCTGAGAGACACCAAAAAGTAAAGAAATTC	1272
Db	565	YMSNNAMWYRRYSARWSSMARWTTTNNNNMMSGVRRMAGTM	523

RESULT 8
 US-10-101-464A-251
 Sequence 251, Application US/10101464A
 Patent No. 6768041
 GENERAL INFORMATION:
 APPLICANT: Strabala, Timothy
 APPLICANT: Nieuwenhuizen, Nicolaas
 APPLICANT: Higgins, Colleen M.
 TITLE OF INVENTION: Compositions Isolated from Plant Cells
 TITLE OR INVENTION: and Their Use in the Modification of Plant Cell Signaling
 PILE REFERENCE: 11000.1020c2
 CURRENT APPLICATION NUMBER: US/10/101.464A
 CURRENT FILING DATE: 2002-03-18
 PRIOR APPLICATION NUMBER: 09/704,302
 PRIOR FILING DATE: 2000-11-01
 PRIOR APPLICATION NUMBER: 09/728,986
 PRIOR FILING DATE: 1999-01-12
 PRIOR APPLICATION NUMBER: 60/162,866
 PRIOR FILING DATE: 1999-11-01
 PRIOR APPLICATION NUMBER: PCT/US00/00724
 PRIOR FILING DATE: 2000-01-11

```

; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 251
; LENGTH: 937
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-10-101-464A-251

Query Match      3.4%; Score 43; DB 4; Length 937;
Best Local Similarity 51.9%; Pred. No. 0.0037;
Matches 97; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY      1060 TATGTCCAAATACCTAGAGAACTATGAAAGATCCGGTCATCGCATGATGCTATTGCAT 1119
DB      732 TGTGCCCCATCCTTCAGGAATATATGATGATCCTTCAGATTGGCGCGATGTTTCACTT 791
QY      1120 ATGAAAGAGAGCAATGGAATAATGTGATCAGCAAAAAGAACTACAGATCCCATGACAA 1179
DB      792 ATGAGAGAGAGCAATGAGCGGAATGTTGACACGACGATACATCGCCAAATGACCA 851
QY      1180 ATCTTGTCTTCTCTCAGCGGTACTTACACCAATAGAGACTTGAATAGGCCATCAATA 1239
DB      852 ACTTGAAGCTTACTGATTCATTTGATCTCATCTCCCAACGACGCTTACGGTGGCAATTCAAG 911
QY      1240 GATGGCT 1246
DB      912 AGTGGCT 918

RESULT 9
US-09-345-882-1/c
; Sequence 1, Application US/09345882
; Patent No. 6399373
; GENERAL INFORMATION:
; APPLICANT: Bougueteloret, Lydie
; TITLE OF INVENTION: A NOCLETIC ACID ENCODING A RETINOBLASTOMA BINDING PROTEIN (RBP-7)
; FILE REFERENCE: GENSET.031A
; CURRENT APPLICATION NUMBER: US/09/345,882
; CURRENT FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: US 60/091,315
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/111,909
; PRIOR FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: Patent.pm
; SEQ ID NO 1
; LENGTH: 162450
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 72794 : polymorphic base A or G
; OTHER INFORMATION: 5-124-273 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 88073 : polymorphic base A or C
; OTHER INFORMATION: 5-127-261 : polymorphic base A or C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 90842 : polymorphic base A or G
; OTHER INFORMATION: 99-1437-325 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 97122 : polymorphic base deletion of GT
; OTHER INFORMATION: 99-1442-224 : polymorphic base G or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 97152 : polymorphic base G or T

; OTHER INFORMATION: 5-129-144 : polymorphic base deletion of T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 99098 : polymorphic base A or G
; OTHER INFORMATION: 5-130-257 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 99117 : polymorphic base A or G
; OTHER INFORMATION: 5-130-276 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 103805 : polymorphic base A or T
; OTHER INFORMATION: 5-131-395 : polymorphic base A or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 106940 : polymorphic base insertion of A
; OTHER INFORMATION: 5-133-375 : polymorphic base insertion of A
; FEATURE:
; NAME/KEY: allele
; LOCATION: 108106 : polymorphic base insertion of A
; OTHER INFORMATION: 5-135-155 : polymorphic base insertion of A
; FEATURE:
; NAME/KEY: allele
; LOCATION: 108149 : polymorphic base insertion of GTTT
; OTHER INFORMATION: 5-135-198 : polymorphic base insertion of GTTT
; FEATURE:
; NAME/KEY: allele
; LOCATION: 108308 : polymorphic base A or G
; OTHER INFORMATION: 5-135-357 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 108471 : polymorphic base C or T
; OTHER INFORMATION: 5-136-174 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 134134 : polymorphic base C or T
; OTHER INFORMATION: 5-140-120 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 134362 : polymorphic base insertion of A
; OTHER INFORMATION: 5-140-348 : polymorphic base insertion of A
; FEATURE:
; NAME/KEY: allele
; LOCATION: 134374 : polymorphic base insertion of CA
; OTHER INFORMATION: 5-140-361 : polymorphic base insertion of CA
; FEATURE:
; NAME/KEY: allele
; LOCATION: 146328 : polymorphic base A or G
; OTHER INFORMATION: 5-143-84 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 146345 : polymorphic base A or C
; OTHER INFORMATION: 5-143-101 : polymorphic base A or C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 150329 : polymorphic base A or G
; OTHER INFORMATION: 5-145-24 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 160031 : polymorphic base G or T
; OTHER INFORMATION: 5-148-352 : polymorphic base G or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 72771..72817 : polymorphic fragment 5-124-273 SEQ ID30
; OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID30
; FEATURE:
; NAME/KEY: allele
; LOCATION: 72771..72817 : polymorphic fragment 5-124-273 SEQ ID51
; OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID51
; FEATURE:
; NAME/KEY: allele
; LOCATION: 88050..88096 : polymorphic fragment 5-127-261 SEQ ID31
; OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID31
```

	FEATURE:	
	NAME/KEY: allele	
	LOCATION: 88050..88096	
	OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID52	
	FEATURE:	
	NAME/KEY: allele	
	LOCATION: 90819..90865	
	OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID49	
	FEATURE:	
	NAME/KEY: allele	
	LOCATION: 90819..90865	
	OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID70	
	FEATURE:	
	NAME/KEY: allele	
	LOCATION: 93690..93736	
	OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID32	
	FEATURE:	
	NAME/KEY: allele	
	LOCATION: 93690..93736	
	OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID53	
	FEATURE:	
	NAME/KEY: allele	
	LOCATION: 97099..97145	
	OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID50	
	FEATURE:	
	NAME/KEY: allele	
	LOCATION: 97099..97145	
	OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID71	
	FEATURE:	
	NAME/KEY: allele	
	LOCATION: 97130..97177	
	OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID33	
	FEATURE:	
	NAME/KEY: allele	
	LOCATION: 97130..97177	
	OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID54	
	FEATURE:	
	NAME/KEY: allele	
	LOCATION: 99075..99121	
	OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID34	
	FEATURE:	
	NAME/KEY: allele	
	LOCATION: 99075..99121	
	OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID55	
	FEATURE:	
	NAME/KEY: allele	
	LOCATION: 99094..99140	
	OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID35	
	FEATURE:	
	NAME/KEY: allele	
	LOCATION: 99094..99140	
	OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID56	
	FEATURE:	
	NAME/KEY: allele	
	LOCATION: 103783..103828	
	OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID36	
	FEATURE:	
	NAME/KEY: allele	
	LOCATION: 103783..103828	
	OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID57	
	FEATURE:	
	NAME/KEY: allele	
	LOCATION: 106918..106966	
	OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID37	
	FEATURE:	
	NAME/KEY: allele	
	LOCATION: 106918..106966	
	OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID58	
	FEATURE:	
	NAME/KEY: allele	
	LOCATION: 108084..108130	
	OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID38	
	FEATURE:	

```

NAME/KEY allele
LOCATION: 108084..108130
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID59
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID39
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID60
FEATURE:

Query Match 3.2%; Score 40.4; DB 3; Length 162450;
Best Local Similarity 52.4%; Pred. No. 1;
Matches 89; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 1021 AGGTAAATCCCTTCTTCAGGAATCCGATGAATTTATATGTCATATACATAGAAC 1080
DB 73113 AGAGAAATATATTATTTAACTGAGTAGTAAACAGTTATATCCAAAATATATTAAGAAC 73655
QY 1081 TTATGAAAGATCCGGTCATCGATCGATGCGTATTCATATATGAAAAAGAACATGGAAA 1140
DB 73653 ACATATACCAAACTAATAAGAAAAAGATTAATCATCTCAAGAAAAAGGAAAGAAAT 73599
QY 1141 ATTGATATCGCAAAAAGAAAGCTACAGTCCCATGACAAATCTGTCTT 1190
DB 73593 AAAGACCCAGCACTCAAAAACGAGAAAAACAAATGAATATATGTTCTT 73544

RESULT 10
US-08-188-582-17
Sequence 17, Application US/08188582
Patent No. 5534410
GENERAL INFORMATION:
APPLICANT: Tjian, Robert
APPLICANT: Comai, Lucio
APPLICANT: Dynlacht, Brian D.
APPLICANT: Hoey, Timothy
APPLICANT: Ruppert, Siegfried
APPLICANT: Tanese, Naoko
APPLICANT: Wang, Edith
APPLICANT: Weinzierl, Robert O.J.
TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAFs AND METHODS OF USE
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHER, HOHBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/188,582
FILING DATE: 28-JAN-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 2152 base pairs

```

TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2112
US-08-188-582-17

Query Match 3.1%; Score 40; DB 1; Length 2152;
Best Local Similarity 48.3%; Pred. No. 0.065;
Matches 112; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY 407 TGTGAGGAGTTTGGCAGTTTCCCGACGCTCCAGCTGTTTGGCATCAGGGGCGAGCTGATG 466
DB 1599 TGTGAATTGTACAGAGATTCATCCAAATTCTAATTATGTTGCTACGGGCTCTGCAGACAG 1658
QY 467 AACTGTGTTTGTGGAATGCACAGTCATCAAAATTATATAGATGTGTATGAAGA 526
DB 1659 AACTGTGCGGCTCTGGGACGCTCTGAATGGTAACTGTGTAAGATCTTCACTGACACAA 1718
QY 527 TGGCTCCTTGGGGCATGTGCATTTCTCTAATGGAAGCTCTTTGTCACTGGCTCTTC 586
DB 1719 GGGACCAATTCATCTCTTGACATTTTCTCCCAATGGAGATTCCTGCTACAGAGCAAC 1778
QY 587 ATGTGTGATTTTAACAGTGTGGATGATTAATGAGGTGTCTGCATAGTGA 638
DB 1779 AGATGGCAGAGTGTCTTCTTGGGATATGTGACATGTTGATGTGGAGAA 1830

RESULT 11
US-08-646-715-17
Sequence 17, Application US/08646715
Patent No. 5637686

GENERAL INFORMATION:
APPLICANT: Tjian, Robert
APPLICANT: Comai, Lucio
APPLICANT: Dynlacht, Brian D.
APPLICANT: Hoey, Timothy
APPLICANT: Ruppert, Siegfried
APPLICANT: Tanese, Naoko
APPLICANT: Wang, Edith
APPLICANT: Weinzierl, Robert O.J.
TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAFs AND METHODS OF USE
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOMBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/646,715
FILING DATE: 09-MAY-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/188,582
FILING DATE: 28-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-57650-2/ALT/BAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 2152 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2112
US-08-646-715-17

Query Match 3.1%; Score 40; DB 1; Length 2152;
Best Local Similarity 48.3%; Pred. No. 0.065;
Matches 112; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY 407 TGTGAGGAGTTTGGCAGTTTCCCGACGCTCCAGCTGTTTGGCATCAGGGGCGAGCTGATG 466
DB 1599 TGTGAATTGTACAGATTCATCCAAATTCTAATTATGTTGCTACGGGCTCTGCAGACAG 1658
QY 467 AACTGTGTTTGTGGAATGCACAGTCATCAAAATTATATAGATGTGTATGAAGA 526
DB 1659 AACTGTGCGGCTCTGGGACGCTCTGAATGGTAACTGTGTAAGATCTTCACTGACACAA 1718
QY 527 TGGCTCCTTGGGGCATGTGCATTTCTCTAATGGAAGCTCTTTGTCACTGGCTCTTC 586
DB 1719 GGGACCAATTCATCTCTTGACATTTTCTCCCAATGGAGATTCCTGCTACAGAGCAAC 1778
QY 587 ATGTGTGATTTTAACAGTGTGGATGATTAATGAGGTGTCTGCATAGTGA 638
DB 1779 AGATGGCAGAGTGTCTTCTTGGGATATGTGACATGTTGATGTGGAGAA 1830

RESULT 12
US-08-724-394A-20
Sequence 20, Application US/08724394A
Patent No. 5872237

GENERAL INFORMATION:
APPLICANT: Feder, John N.
APPLICANT: Kronmal, Gregory S.
APPLICANT: Lauer, Peter M.
APPLICANT: Ruddy, David A.
APPLICANT: Thomas, Winston
APPLICANT: Tsuchihashi, Zenta
APPLICANT: Wolff, Roger K.
TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1
TITLE OF INVENTION: Sequences and Antibodies thereo
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-9834

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/724,394A
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Filtz, Renee A.
REGISTRATION NUMBER: 35,136
REFERENCE/DOCKET NUMBER: 017957-000100
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:

```

;
; LENGTH: 246240 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..246240
; OTHER INFORMATION: /note= "HLA-H.CONTIG"
US-08-724-394A-20

Query Match
Best Local Similarity 55.9%; Score 40; DB 2; Length 246240;
Matches 76; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 773 TACCATATCTTACCAAGCGCCACAGAACATCAGCTGAGAGCAATTTACCGAAGTTGCTC 832
DB 45039 TACTCATGCGCTGATGAGTATTACCAACATACCCCTGGGCCAATTAACAAACAGCAGAGC 45098

QY 833 AGAGAGGTGCTCTCAACATGCTTTGTGCACAGATTTAAAGATCTTGTGTAATTT 892
DB 45099 AAATGTGTTGTGTGTATTAATCTTTCTACACAAATACAGAAAAAGTATTCATGTT 45158

QY 893 CAAGATGATTAACATT 908
DB 45159 CAGCATTAATTAATT 45174

RESULT 13
US-08-724-394A-21
; Sequence 21, Application US/08724394A
; Patent No. 5872237
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Krommal, Gregory S.
; APPLICANT: Laufer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1
; TITLE OF INVENTION: Sequences and Antibodies Thereo
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,394A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Flits, Renee A.
; REGISTRATION NUMBER: 35,136
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 246240 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: cdna
```

```

;
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..246240
; OTHER INFORMATION: /note= "HLA-H.CONTIG"
US-08-724-394A-21

Query Match
Best Local Similarity 55.9%; Score 40; DB 2; Length 246240;
Matches 76; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 773 TACCATATCTTACCAAGCGCCACAGAACATCAGCTGAGAGCAATTTACCGAAGTTGCTC 832
DB 45039 TACTCATGCGCTGATGAGTATTACCAACATACCCCTGGGCCAATTAACAAACAGCAGAGC 45098

QY 833 AGAGAGGTGCTCTCAACATGCTTTGTGCACAGATTTAAAGATCTTGTGTAATTT 892
DB 45099 AAATGTGTTGTGTGTATTAATCTTTCTACACAAATACAGAAAAAGTATTCATGTT 45158

QY 893 CAAGATGATTAACATT 908
DB 45159 CAGCATTAATTAATT 45174

RESULT 14
US-08-724-394A-22
; Sequence 22, Application US/08724394A
; Patent No. 5872237
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Krommal, Gregory S.
; APPLICANT: Laufer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1
; TITLE OF INVENTION: Sequences and Antibodies Thereo
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,394A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Flits, Renee A.
; REGISTRATION NUMBER: 35,136
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 246240 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..246240
; OTHER INFORMATION: /note= "HLA-H.CONTIG"
US-08-724-394A-22
```

Query Match	3.1%	Score 40;	DB 2;	Length 246240;
Best Local Similarity	55.9%;	Pred. No. 1.9;		
Matches 76;	Conservative 0;	Mismatches 60;	Indels 0;	Gaps 0;

Qy	773	TACCCATATTTTGGCAAGGGGACAGAA	CATCTAGCTGAAGCAATTTACCGAAGATGGTC	832
Db	45039	TACTCATGCTGTATGAAATTTTACCA	AACTATCCCTGGGCCATTAACAAACGAGAGC	450988
Qy	833	AGAGAGAGTGCCTCAACATGCGTTTGTG	CACAGATTTAAAGATCTTGGTATTTT	892
Db	45099	AAATGTGTTGTGTGTATTA	CTTCTACACAAATACAGAAAAAGTGA	451588
Qy	893	CAAGATGAATACATT	908	
Db	45159	CAGCATAAATTAATT	45174	

```

RESULT 15
US-09-680-364-10
; Sequence 10, Application US/09690364
; Patent No. 6468795
; GENERAL INFORMATION:
; APPLICANT: Hong Zhang
; APPLICANT: Andrew T. Watt
; TITLE OF INVENTION: ANTISENSE MODULATION OF APAF-1 EXPRESSION
; FILE REFERENCE: RTS-0190
; CURRENT APPLICATION NUMBER: US/09/690.364
; CURRENT FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 100
; SEQ ID NO 10
; LENGTH: 5152
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (586)...(4302)
; US-09-690-364-10

```

	Query Match	3.1%	Score 39.6;	DB 3,	Length 5152;
	Best Local Similarity	54.9%	Pred. No. 0.16;		
	Matches	78;	Conservative	0;	Mismatches 64;
				Indels	0;
				Gaps	0;
QY	442	GTITGGCATCAGGGGACAGTCATGAACTGTGTTTGTGAAATGCACAGTCATACAAAT	501		
Db	3722	GATTGCTTTCTTGGCATTTGTCATTTGATGAAACGGTAGAGGTGTGATGTCATTAACGGAGAA	3781		
QY	502	TATATAGATGTGTGTAAGTGTAAAGATGCTCTTGGGGGCATGTGCATTTTCTCTATG	561		
Db	3782	TAGAAAGACATTTACTGTCTATCAGGGACAGATGCTTTCTGTGTCTATCTTCTGTAG	3841		
QY	562	GAAGCTTCTTGTCACTGGCTC	583		
Db	3842	CGACCAAGTTTCTCTTACTCTC	3863		

Search completed: February 5, 2005, 08:22:39
Job time : 252 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 5, 2005, 07:02:01 / Search time 5658 Seconds
(without alignments)
10893.431 Million cell updates/sec

Title: US-10-077-111-12

Perfect score: 1272
Sequence: 1 gaattcgcttcacatcg.....acaccaaagtaagaattc 1272

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 2805242

Minimum DB seq length: 0
Maximum DB seq length: 200

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: GenEmbl:
1: gb_ba:*
2: gb_brg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	176.8	13.9	180	6	CQ110706 Sequence
C 2	176.8	13.9	180	6	CQ149443 Sequence
C 3	176.8	13.9	180	6	CQ232723 Sequence
C 4	176.8	13.9	180	6	CQ270754 Sequence
C 5	176.8	13.9	180	6	CQ308036 Sequence
C 6	176.8	13.9	180	6	CQ344856 Sequence
C 7	132	10.4	132	6	CQ058095 Sequence
C 8	132	10.4	132	6	CQ077416 Sequence
C 9	132	10.4	132	6	CQ108429 Sequence
C 10	132	10.4	132	6	CQ147070 Sequence
C 11	132	10.4	132	6	CQ206855 Sequence
C 12	132	10.4	132	6	CQ230294 Sequence
C 13	132	10.4	132	6	CQ268431 Sequence
C 14	132	10.4	132	6	CQ305465 Sequence
C 15	32.6	2.6	157	6	AR248923 Sequence
C 16	32	2.5	181	11	BX000936 Arabidops
C 17	32	2.5	181	11	BX000945 Arabidops
C 18	31	2.4	165	6	CQ704579 Sequence
C 19	29.8	2.3	195	11	G43786 WAF-3389-S

20	29.6	2.3	136	6	CQ110718 Sequence
21	29.6	2.3	136	6	CQ149455 Sequence
22	29.6	2.3	136	6	CQ232735 Sequence
23	29.6	2.3	136	6	CQ270766 Sequence
24	29.6	2.3	136	6	CQ308048 Sequence
25	29.6	2.3	136	6	CQ344868 Sequence
C 26	29.6	2.3	182	4	EMAJ4737
C 27	29.2	2.3	153	1	MC208
C 28	29.2	2.3	171	1	LBW5683
C 29	29.2	2.3	140	6	AX276370
30	28.8	2.3	152	6	CQ077854 Sequence
31	28.8	2.3	152	6	CQ108861 Sequence
32	28.8	2.3	152	6	CQ182858 Sequence
33	28.8	2.3	152	6	CQ207282 Sequence
34	28.8	2.3	153	6	CQ056465 Sequence
35	28.8	2.3	153	6	CQ075732 Sequence
36	28.8	2.3	153	6	CQ106713 Sequence
37	28.8	2.3	153	6	CQ145366 Sequence
38	28.8	2.3	153	6	CQ180805 Sequence
39	28.8	2.3	153	6	CQ205181 Sequence
40	28.8	2.3	153	6	CQ28561 Sequence
41	28.8	2.3	153	6	CQ266716 Sequence
42	28.8	2.3	153	6	CQ303668 Sequence
43	28.8	2.3	153	6	CQ340999 Sequence
44	28.8	2.3	156	10	AF357446 Mus muscu
45	28.6	2.2	150	11	AL935341 Arabidops

ALIGNMENTS

RESULT 1	CQ110706/c	Sequence 19565 from Patent WO0157272.	180 bp	DNA	linear	PAT 21-JAN-2004
LOCUS	CQ110706	Sequence 19565 from Patent WO0157272.				
DEFINITION	CQ110706	Sequence 19565 from Patent WO0157272.				
ACCESSION	CQ110706	Sequence 19565 from Patent WO0157272.				
VERSION	CQ110706.1	GI:41080089				
KEYWORDS						
SOURCE						
ORGANISM	Homo sapiens (human)					
REFERENCE						
AUTHORS	Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.					
TITLE	Human genome-derived single exon nucleic acid probes useful for					
JOURNAL	Analysis of gene expression in human placenta					
FEATURES						
source	1.180	location/Qualifiers				
	/organism="Homo sapiens"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:9606"					
	/note="MAP TO AC006501.5-EXPRESSED IN PLACENTA. SIGNAL =					
	1.9-BST HUMAN HIT: AL040518.1, EVALUE 1.00e-94,-NT HIT:					
	U45139.1, EVALUE 1.10e-01-SWISSPROT HIT: Q16760, EVALUE					
	6.80e-02"					
ORIGIN						
Query Match	13.9%	Score 176.8; DB 6; Length 180;				
Best Local Similarity	98.9%	Pred. No. 1.1e-36;				
Matches 178; Conservative	0; Mismatches 2; Indels 0; Gaps 0;					
QY	787	CAAGGCGCAGCAACATCACTGAGCAATTTTCCAGAGATTGTCAGAGAGCTGCT	846			
Db	180	CAAGGCGCAGCAACATCACTGAGCAATTTTCCAGAGATTGTCAGAGAGAGTCT	121			
QY	847	CAACATGCGCTTTGTGCAACAAGATTAAAGATTTGTTGATTTTCAAGATTAATPACA	906			
Db	120	CAACATGCGCTTTGTGCAACAAGATTAAAGATTTGTTGATTTTCAAGATTAATPACA	61			
QY	907	TTGATGGAAGAAGACTGTGAATCTTACAAAGAGTGGCTGATGATTTGAATTTG	966			

Db	60	TTGATGGAAAAGAACTGTTGAATCTTACAAAGAAAGAGTCGCTGATGATTTGAAAATTG	1
RESULT 2			
LOCUS	CQ149443	180 bp	DNA
DEFINITION	Sequence 19465 from Patent WO0157276.		linear
ACCESSION	CQ149443		
VERSION	CQ149443.1	GI:41156793	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE			
AUTHORS	Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.		
TITLE	Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human bone marrow		
JOURNAL	Patent: WO 0157276-A 19465 09-AUG-2001;		
FEATURES			
source	Location/Qualifiers		
	1..180		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
	/note="MAP TO AC006501.5-EXPRESSED IN BONE MARROW, SIGNAL = 0.81-EST HUMAN HIT: AL040518.1, EVALUE 1.00e-94-NT HIT: U43339.1, EVALUE 1.10e-01-SWISSPROT HIT: Q16760, EVALUE 6.80e-02"		
ORIGIN			
Query Match	13.9%;	Score 176.8;	DB 6; Length 180;
Best Local Similarity	98.9%;	Pred. No. 1.1e-36;	
Matches 178;	Conservative 0;	Mismatches 2;	Indels 0; Gaps 0;
Oy	787	CAAGCGGACACGAATCAGCTGAGCAATTTACCGAAGATTGTCAGAGAGTCGCT	846
Db	180	CAAGCGGACACGAATCAGCTGAGCAATTTACCGAAGATTGTCAGAGAGATGCT	121
Oy	847	CAACATGCGCTTTGTGACACAGATTTTAAAGATCTTGTGATTTTCAAGATGAATA	906
Db	120	CAACATGCGCTTTGTGACACAGATTTTAAAGATCTTGTGATTTTCAAGATGAATA	61
Oy	907	TTGATGAAAAAGAACTGTTGAATCTTACAAAGAAAGAGTCGCTGATGATTTGAAAATTG	966
Db	60	TTGATGAAAAAGAACTGTTGAATCTTACAAAGAAAGAGTCGCTGATGATTTGAAAATTG	1
RESULT 3			
LOCUS	CQ323723	180 bp	DNA
DEFINITION	Sequence 19562 from Patent WO0157273.		linear
ACCESSION	CQ323723		
VERSION	CQ323723.1	GI:41215941	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE			
AUTHORS	Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.		
TITLE	HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER<130> PB 0004 WO		
JOURNAL	Patent: WO 0157273-A 19562 09-AUG-2001;		

Query Match	13.9%	Score 176.8	DB 6	Length 180
Best Local Similarity	98.9%	Pred. No. 1.1e-36		
Matches 178	Conservative 0	Mismatches 2	Indels 0	Gaps 0
Db	787	CAAGGCCACGAACATCAGCTGAGAGCAATTTACCGAAGATTGGTCAGAGAGGTCTCT	846	
Db	180	CAAGGCCACGAACATCAGCTGAGAGCAATTTACCGAAGATTGGTCAGAGAGGTCTCT	121	
Qy	847	CAACATGCTTTGGCACAAGATTAAAGATCTTGTGGTATTTTCAAGTGAATACAA	906	
Db	120	CAACATGCTTTGGCACAAGATTAAAGATCTTGTGGTATTTTCAAGTGAATACAA	61	
Qy	907	TTGATGAGAAAGAACTGTGATCTTCAAAAGAAAGCTGGCTGATGATTTGAAAATTG	966	
Db	60	TTGATGAGAAAGAACTGTGATCTTCAAAAGAAAGCTGGCTGATGATTTGAAAATTG	1	
RESULT 4				
LOCUS	CO270754/c	180 bp	DNA	linear PAT 23-JAN-2004
DEFINITION	Sequence 19015 from Patent WO0157277.			
ACCESSION	CO270754			
VERSION	CO270754.1	GI:41243358		
KEYWORDS				
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.			
AUTHORS	Human genome-derived single exon nucleic acid probes useful for			
TITLE	analysis of gene expression in human fetal liver			
JOURNAL	Patent: WO 0157277-A 19015 09-AUG-2001;			
	Aeomica, Inc. (US)			
FEATURES	Location/Qualifiers			
source	1..180			
	/organism="Homo sapiens"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:9606"			
	/note="MAP TO AC006501.5-EXPRESSED IN FETAL LIVER. SIGNAL			
	= 1.3-E8T HUMAN HIT: AL040518.1, EVALUE 1.00e-94-NT HIT:			
	U41339.1, EVALUE 1.10e-01-SWISSPROT HIT: Q16760, EVALUE			
	6.80e-02"			
ORIGIN				
Query Match	13.9%	Score 176.8	DB 6	Length 180
Best Local Similarity	98.9%	Pred. No. 1.1e-36		
Matches 178	Conservative 0	Mismatches 2	Indels 0	Gaps 0
Db	787	CAAGGCCACGAACATCAGCTGAGAGCAATTTACCGAAGATTGGTCAGAGAGGTCTCT	846	
Db	180	CAAGGCCACGAACATCAGCTGAGAGCAATTTACCGAAGATTGGTCAGAGAGGTCTCT	121	
Qy	847	CAACATGCTTTGGCACAAGATTAAAGATCTTGTGGTATTTTCAAGTGAATACAA	906	
Db	120	CAACATGCTTTGGCACAAGATTAAAGATCTTGTGGTATTTTCAAGTGAATACAA	61	
Qy	907	TTGATGAGAAAGAACTGTGATCTTCAAAAGAAAGCTGGCTGATGATTTGAAAATTG	966	
Db	60	TTGATGAGAAAGAACTGTGATCTTCAAAAGAAAGCTGGCTGATGATTTGAAAATTG	1	
RESULT 5				

CQ308036/c
LOCUS CQ308036 180 bp DNA linear PAT 23-JAN-2004
DEFINITION Sequence 19141 from Patent WO0166003.
ACCESSION CQ308036
VERSION CQ308036.1 GI:41268613
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Chordata; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
Human genome-derived single exon nucleic acid probes useful for
analysis of gene expression in human lung
Patent: WO 0166003-A 19141 15-NOV-2001;
Aeomica, Inc. (US)
FEATURES
Source
1. 180
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC006501.5-EXPRESSED IN LUNG, SIGNAL =
0.88-BEST_HUMAN HIT: AL040518.1, EVALUE 1.00e-94-NT HIT:
U43139.1, EVALUE 1.10e-01-SWISSPROT HIT: Q16760, EVALUE
6.80e-02"
ORIGIN
Query Match 13.9%; Score 176.8; DB 6; Length 180;
Best Local Similarity 98.9%; Pred. No. 1.1e-36;
Matches 178; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 787 CAAGGCGCACAGAACATCAGCTGAAGCAATTTACGAAGATTGGTCAGAGAGTGTCT 846
180 CAAGGCGCACAGAACATCAGCTGAAGCAATTTACGAAGATTGGTCAGAGAGATGTCT 121
DB 847 CAACATGGCTTTGTGACCAAGATTTAAAGATCTTGTGATTTTCAAGATGATAC 906
120 CAACATGGCTTTGTGACCAAGATTTAAAGATCTTGTGATTTTCAAGATGATAC 61
QY 907 TTGATGAAAAAGAACTGTTGAATTTACAAAGAAAGTGGCTGATGATTTGAAAAATG 966
60 TTGATGAAAAAGAACTGTTGAATTTACAAAGAAAGTGGCTGATGATTTGAAAAATG 1
DB
RESULT 6
CQ344856/c
LOCUS CQ344856 180 bp DNA linear PAT 23-JAN-2004
DEFINITION Sequence 18950 from Patent WO0157275.
ACCESSION CQ344856
VERSION CQ344856.1 GI:41293927
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Chordata; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
Human genome-derived single exon nucleic acid probes useful for
analysis of gene expression in human brain
Patent: WO 0157275-A 18950 09-AUG-2001;
Aeomica, Inc. (US)
FEATURES
Source
1. 180
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC006501.5-EXPRESSED IN BRAIN, SIGNAL =
0.73-BEST_HUMAN HIT: AL040518.1, EVALUE 1.00e-94-NT HIT:
U43139.1, EVALUE 1.10e-01-SWISSPROT HIT: Q16760, EVALUE
6.80e-02"
ORIGIN
Query Match 13.9%; Score 176.8; DB 6; Length 180;

Best Local Similarity 98.9%; Pred. No. 1.1e-36;
Matches 178; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 787 CAAGGCGCACAGAACATCAGCTGAAGCAATTTACGAAGATTGGTCAGAGAGTGTCT 846
180 CAAGGCGCACAGAACATCAGCTGAAGCAATTTACGAAGATTGGTCAGAGAGATGTCT 121
DB 847 CAACATGGCTTTGTGACCAAGATTTAAAGATCTTGTGATTTTCAAGATGATAC 906
120 CAACATGGCTTTGTGACCAAGATTTAAAGATCTTGTGATTTTCAAGATGATAC 61
QY 907 TTGATGAAAAAGAACTGTTGAATTTACAAAGAAAGTGGCTGATGATTTGAAAAATG 966
60 TTGATGAAAAAGAACTGTTGAATTTACAAAGAAAGTGGCTGATGATTTGAAAAATG 1
DB
RESULT 7
CQ058095/c
LOCUS CQ058095 132 bp DNA linear PAT 19-JAN-2004
DEFINITION Sequence 8915 from Patent WO0157270.
ACCESSION CQ058095
VERSION CQ058095.1 GI:41032601
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Chordata; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
Human genome-derived single exon nucleic acid probes useful for
analysis of gene expression in human breast and hbl 100 cells
Patent: WO 0157270-A 8915 09-AUG-2001;
Aeomica, Inc. (US)
FEATURES
Source
1. 132
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC009307.1-EXPRESSED IN HBL100, SIGNAL =
0.68-BEST_HUMAN HIT: BB675766.1, EVALUE 2.00e-68-NT HIT:
X71133.1, EVALUE 1.20e+00-SWISSPROT HIT: P54860, EVALUE
1.00e+00"
ORIGIN
Query Match 10.4%; Score 132; DB 6; Length 132;
Best Local Similarity 100.0%; Pred. No. 1.6e-24;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 976 GACTGGGTAGTAAGCTGAGGAAATGAAAGAGCTCAGACCAAGTTAAATCCCTTT 1035
132 GACTGGGTAGTAAGCTGAGGAAATGAAAGAGCTCAGACCAAGTTAAATCCCTTT 73
DB 1036 CTTCAGGAATTCCTGATGAAATTTATATGTCATACTAGAGAACTTATGAAAGATCCGG 1095
72 CTTCAGGAATTCCTGATGAAATTTATATGTCATACTAGAGAACTTATGAAAGATCCGG 13
QY 1096 TCATGCGATCAG 1107
12 TCATGCGATCAG 1
DB
RESULT 8
CQ077416/c
LOCUS CQ077416 132 bp DNA linear PAT 20-JAN-2004
DEFINITION Sequence 13216 from Patent WO0157278.
ACCESSION CQ077416
VERSION CQ077416.1 GI:41047285
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Chordata; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1

AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
TITLE Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human hela cells or other human cervical epithelial cells
JOURNAL Patent: WO 0157278-A 13216 09-AUG-2001;
Aeomica, Inc. (US)

ORIGIN

Query Match	10.4%	Score 132;	DB 6;	Length 132;
Best Local Similarity	100.0%	Pred. No. 1.6e-24;		
Matches 132;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

9y 977 GACTGCGTGTAAAGTGTCTGAGGAAATTTGAAGAGCTCAGGACCAAGTTAAATCCCTTT 1035

Db 132 GACTGCGTGTAAAGTGTCTGAGGAAATTTGAAGAGCTCAGGACCAAGTTAAATCCCTTT 73

QY 1036 CTTGAGGAATTCCTGATGAATTATATGTCCAATCACTGAGAACTTATGAAGAATCCGG 1095

Db 72 CTTGAGGAATTCCTGATGAATTATATGTCCAATCACTGAGAACTTATGAAGAATCCGG 13

Qy	1096	TCATCGCATCAG	1107
Db	12	TCATCGCATCAG	1

ORIGIN

RESULT 9			
CQ108429/c			
LOCUS	CQ108429	132 bp	DNA
DEFINITION	Sequence 17288 from Patent WO0157272.		linear PAT 21-JAN-2004

SOURCE ORGANISM	Source
Homo sapiens (human)	Human
Homo sapiens	Human

REFERENCE
AUTHORS 1 Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
TITLE Human genome-derived single exon nucleic acid probes useful for
JOURNAL analysis of gene expression in human placenta
Patent: WO 0157272-A 17288 09-AUG-2001;
Aecmca, Inc. (US)

ORIGIN

Query Match	10.4%	Score 132;	DB 6;	Length 132;
Best Local Similarity	100.0%	Pred. No. 1.6e-24;		
Matches 132; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	976	GAC	T	C	G	T	A	A	G	C	T	G	A	G	A	A	A	T	T	G	A	G	C	T	C	A	G	A	C	C	A	G	T	T	A	A	T	C	C	T	1035		
Db	132	G	A	C	T	C	G	T	A	A	G	C	T	G	A	G	A	A	T	T	G	A	G	A	C	T	C	A	G	A	C	C	A	G	T	T	A	A	T	C	C	T	73

Qy 1036 CTTAGGAAATTCCTGATGATTTTATATGTCCAACTACTAGGAACTTATGANAAGTCCGG 1035

Db 72 CTTAGGAAATTCCTGATGATTTTATATGTCCAACTACTAGGAACTTATGANAAGTCCGG 13

QY	1096	TCATCGCATCAG	1107
Db	12	TCATCGCATCAG	1

RESULT 10			
COL47070/c			
LOCUS	COL47070	132 bp	DNA
DEFINITION	Sequence 17092 from Patent WO0157276.		linear
ACCESSION	COL47070		
VERSION	COL47070.1	GI:41154420	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
	Homo sapiens		

SOURCE	ORGANISM
Homo sapiens (human)	Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	

REFERENCE
AUTHORS
TITLE
JOURNAL
1 Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
Human genome-derived single exon nucleic acid probes useful for
analysis of gene expression in human bone marrow
Parent: WO 0157276-A 17092 09-ANG-2001;

```

/notes="MAP TO ACC09307.1-EXPRESSIONED IN BONE MARROW, SIGNAL
= 0.75-EST HUMAN HIT: BB675766.1, EVALUE 2.00e-68-NT HIT:
X71133.1, EVALUE 1.20e+00-SWISSPROT HIT: P54860, EVALUE
1.00e+00"

```

```

Query Match      10.4%; Score 132; DB 6; Length 132;
Best Local Similarity 100.0%; Pred. No. 1.6e-24;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

1036 CTTGAGGATTCCTGATGATTTATATGTCCATACTAGAGACTTATGAAAGATCCGG 1095

OY		1096 TCATCGCATCAG	1107
Dd		12 TCATCGCATCAG	1

```
source 1..132
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
```

/note="MAP TO AC009307.1-EXPRESSED IN BT474, SIGNAL =
0.7-EST HUMAN HIT: BE675766.1, EVALUE 2.00e-68-NT HIT:
X71133.1, EVALUE 1.20e+00-SWISSPROT HIT: P54860, EVALUE
1.00e+00"

ORIGIN

Query Match 10.4%; Score 132; DB 6; Length 132;
Best Local Similarity 100.0%; Pred. No. 1.6e-24;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 976 GACTGCGTAGTAAGTCTGAGAAAATTGGAAGCTCAGACCAAGGTTAAATCCCTTT 1035
DB 132 GACTGCGTAGTAAGTCTGAGAAAATTGGAAGCTCAGACCAAGGTTAAATCCCTTT 73
QY 1036 CTTGAGGAATTCCTGAGAAATTTATATGTCCAACTACAGAACTTATGAAAGATCCGG 1095
DB 72 CTTGAGGAATTCCTGAGAAATTTATATGTCCAACTACAGAACTTATGAAAGATCCGG 13
QY 1096 TCATCGCATCAG 1107
DB 12 TCATCGCATCAG 1

RESULT 12
CO30294/c 132 bp DNA linear PAT 21-JAN-2004
LOCUS Sequence 17133 from Patent WO0157273.
DEFINITION CO30294
ACCESSION CO30294
VERSION CO30294.1 GI:41213512
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
TITLE HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER<130> PB 0004 WO

3<150> US 60/180,312<151> 04 February 2000 (04.02.00)<150> US
60/207,456<151> 26 May 2000 (26.05.00)<150> US 09/632,366<151> 03
August 2000 (03.08.00)<150> GB 24263.6<151> 03 October 2000
(03.10.00)<150> US 60/236,359<151> 27 September 2000
(27.09.00)<150> US 60/234,687<151> 21 September 2000
(21.09.00)<150> US 09/608,408<151> 30 June 2000 (30.06.00)<170>
Molecular Dynamics Sequence Listing Engine
Patent: WO 0157273-A 17133 09-AUG-2001;
Aeomica, Inc. (US)

JOURNAL Location/Qualifiers
FEATURES
SOURCE 1..132

/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC009307.1-EXPRESSED IN ADULT LIVER, SIGNAL
= 0.73-EST HUMAN HIT: BE675766.1, EVALUE 2.00e-68-NT HIT:
X71133.1, EVALUE 1.20e+00-SWISSPROT HIT: P54860, EVALUE
1.00e+00"

ORIGIN

Query Match 10.4%; Score 132; DB 6; Length 132;
Best Local Similarity 100.0%; Pred. No. 1.6e-24;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 976 GACTGCGTAGTAAGTCTGAGAAAATTGGAAGCTCAGACCAAGGTTAAATCCCTTT 1035
DB 132 GACTGCGTAGTAAGTCTGAGAAAATTGGAAGCTCAGACCAAGGTTAAATCCCTTT 73
QY 1036 CTTGAGGAATTCCTGAGAAATTTATATGTCCAACTACAGAACTTATGAAAGATCCGG 1095
DB 72 CTTGAGGAATTCCTGAGAAATTTATATGTCCAACTACAGAACTTATGAAAGATCCGG 13
QY 1096 TCATCGCATCAG 1107
DB 12 TCATCGCATCAG 1

RESULT 13

CO268431/c 132 bp DNA linear PAT 23-JAN-2004
LOCUS Sequence 16692 from Patent WO0157277.
DEFINITION CO268431
ACCESSION CO268431
VERSION CO268431.1 GI:41241035
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
TITLE HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR ANALYSIS OF GENE EXPRESSION IN HUMAN FETAL LIVER
Patent: WO 0157277-A 16692 09-AUG-2001;
Aeomica, Inc. (US)

FEATURES
SOURCE 1..132

/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC009307.1-EXPRESSED IN FETAL LIVER, SIGNAL
= 0.78-EST HUMAN HIT: BE675766.1, EVALUE 2.00e-68-NT HIT:
X71133.1, EVALUE 1.20e+00-SWISSPROT HIT: P54860, EVALUE
1.00e+00"

ORIGIN

Query Match 10.4%; Score 132; DB 6; Length 132;
Best Local Similarity 100.0%; Pred. No. 1.6e-24;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 976 GACTGCGTAGTAAGTCTGAGAAAATTGGAAGCTCAGACCAAGGTTAAATCCCTTT 1035
DB 132 GACTGCGTAGTAAGTCTGAGAAAATTGGAAGCTCAGACCAAGGTTAAATCCCTTT 73
QY 1036 CTTGAGGAATTCCTGAGAAATTTATATGTCCAACTACAGAACTTATGAAAGATCCGG 1095
DB 72 CTTGAGGAATTCCTGAGAAATTTATATGTCCAACTACAGAACTTATGAAAGATCCGG 13
QY 1096 TCATCGCATCAG 1107
DB 12 TCATCGCATCAG 1

RESULT 14

CO305465/c 132 bp DNA linear PAT 23-JAN-2004
LOCUS Sequence 16570 from Patent WO0166003.
DEFINITION CO305465
ACCESSION CO305465
VERSION CO305465.1 GI:41266042
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
TITLE HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR ANALYSIS OF GENE EXPRESSION IN HUMAN LUNG
Patent: WO 0166003-A 16570 15-NOV-2001;
Aeomica, Inc. (US)

FEATURES
SOURCE 1..132

/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC009307.1-EXPRESSED IN LUNG, SIGNAL =
0.76-EST HUMAN HIT: BE675766.1, EVALUE 2.00e-68-NT HIT:
X71133.1, EVALUE 1.20e+00-SWISSPROT HIT: P54860, EVALUE
1.00e+00"

ORIGIN

Query Match 10.4%; Score 132; DB 6; Length 132;
 Best Local Similarity 100.0%; Pred. No. 1.6e-24;
 Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 976 GACTGCGCTAGTAAGTCTGAGAAATTTGAAGAGCTCAGAGCCAGCCAAAGTTAAATCCCTTT 1035
 |||
 DB 132 GACTGCGCTAGTAAGTCTGAGAAATTTGAAGAGCTCAGAGCCAGCCAAAGTTAAATCCCTTT 73

QY 1036 CTTCAGGAATTCCTGATGATTTTATATGTCATTAAGTATAGAGAACTTATGAGAAAGATCCGG 1095
 |||
 DB 72 CTTCAGGAATTCCTGATGATTTTATATGTCATTAAGTATAGAGAACTTATGAGAAAGATCCGG 13

QY 1096 TCATCGCATCAG 1107
 |||
 DB 12 TCATCGCATCAG 1

RESULT 15
 AR248923/c 157 bp DNA linear PAT 20-DEC-2002

LOCUS AR248923
 DEFINITION Sequence 4282 from patent US 6476212.
 ACCESSION AR248923
 VERSION AR248923.1 GI:27296797

KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 Unclassified.

REFERENCE 1 (bases 1 to 157)
 AUTHORS Laligudi,R.V., Ito,L.Y. and Sherman,B.K.
 TITL Polynucleotides and polypeptides derived from corn ear
 JOURNAL Patent: US 6476212-A 4282 05-NOV-2002;
 FEATURES Location/Qualifiers

source 1..157

/organism="unknown"
 /mol_type="genomic DNA"

ORIGIN

Query Match 2.6%; Score 32.6; DB 6; Length 157;
 Best Local Similarity 50.0%; Pred. No. 1.6e+03;
 Matches 71; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 66 CGCGCGCGCCGCTCTGCAAGGCTGTTTCTTCAATAAAGACATGTAAGAACTGATT 125
 |||
 DB 149 CGTGCCTTCTGCTCTGAGGCGAGTGATGANNANGCAACGACAGANAAATTGCGAAC 90

QY 126 CACACATTAGCTGATCATGTGACGATGCACTGCTGCTTCTTCCCTTTG 185
 |||
 DB 89 ATCCCAATTGGCNAACCGCNAATGATNGAACAAACTCTAGTGCCTTGTGTNTCCCTCTCA 30

QY 186 GCTACTTGTCTCTTGACAAAA 207
 |||
 DB 29 GGAACGCGGAACCTGTCACAA 8

Search completed: February 5, 2005, 10:22:19
 Job time : 5661 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 5, 2005, 05:27:56 ; Search time 757 Seconds
(without alignments)
9947.042 Million cell updates/sec

Title: US-10-077-111-12

Perfect score: 1272
Sequence: 1 gaattcgcgttcacatcg.....acaccaaagtaagaattc 1272

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 4887668

Minimum DB seq length: 0
Maximum DB seq length: 200

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1:	geneseqn1980s:*
2:	geneseqn1990s:*
3:	geneseqn2000s:*
4:	geneseqn2001as:*
5:	geneseqn2001bs:*
6:	geneseqn2002as:*
7:	geneseqn2002bs:*
8:	geneseqn2003as:*
9:	geneseqn2003bs:*
10:	geneseqn2003cs:*
11:	geneseqn2003ds:*
12:	geneseqn2004as:*
13:	geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	176.8	13.9	180	4	ABA70710 Human foe
2	176.8	13.9	180	4	AAI50879 Probe #19
3	176.8	13.9	180	4	AAI50879 Probe #19
4	176.8	13.9	180	4	AAI50879 Probe #19
5	176.8	13.9	180	4	AAI50879 Probe #19
6	176.8	13.9	180	4	AAI50879 Probe #19
7	176.8	13.9	180	4	AAI50879 Probe #19
8	176.8	13.9	180	4	AAI50879 Probe #19
9	176.8	13.9	180	4	AAI50879 Probe #19
10	176.8	13.9	180	4	AAI50879 Probe #19
11	176.8	13.9	180	4	AAI50879 Probe #19
12	176.8	13.9	180	4	AAI50879 Probe #19
13	176.8	13.9	180	4	AAI50879 Probe #19
14	176.8	13.9	180	4	AAI50879 Probe #19
15	176.8	13.9	180	4	AAI50879 Probe #19
16	176.8	13.9	180	4	AAI50879 Probe #19
17	176.8	13.9	180	4	AAI50879 Probe #19
18	176.8	13.9	180	4	AAI50879 Probe #19
19	176.8	13.9	180	4	AAI50879 Probe #19
20	176.8	13.9	180	4	AAI50879 Probe #19

21	29.6	2.3	136	4	AAI50891	AAI50891 Probe #19
22	29.6	2.3	136	4	AAK44920	AAK44920 Human bon
23	29.6	2.3	136	4	AAK18971	AAK18971 Human bon
24	29.6	2.3	136	4	ABSA4584	ABSA4584 Human liv
25	29.6	2.3	136	4	ABSI19162	ABSI19162 Human gen
26	29.2	2.3	195	2	AAH86637	AAH86637 Human bin
27	29	2.3	180	4	ABA89218	ABA89218 Escherich
28	28.8	2.3	152	4	AAI23721	AAI23721 Probe #13
29	28.8	2.3	152	4	AAI49034	AAI49034 Probe #17
30	28.8	2.3	152	4	ABA50864	ABA50864 Human bre
31	28.8	2.3	152	4	ABA35788	ABA35788 Probe #14
32	28.8	2.3	153	4	AAI21599	AAI21599 Probe #11
33	28.8	2.3	153	4	ABA66672	ABA66672 Human foe
34	28.8	2.3	153	4	AAI46886	AAI46886 Probe #15
35	28.8	2.3	153	4	ABA48763	ABA48763 Human bre
36	28.8	2.3	153	4	ABA33735	ABA33735 Probe #12
37	28.8	2.3	153	4	AAK40831	AAK40831 Human bon
38	28.8	2.3	153	4	AAK15102	AAK15102 Human bra
39	28.8	2.3	153	4	ABSA40410	ABSA40410 Human liv
40	28.8	2.3	153	5	AAI07294	AAI07294 Probe #72
41	28.8	2.3	153	6	ABSI4782	ABSI4782 Human gen
42	28.4	2.2	60	6	ABN39870	ABN39870 Human spl
43	28.2	2.2	194	3	AAK27087	AAK27087 Human sec
44	28	2.2	147	5	ADL37494	ADL37494 Human ova
45	28	2.2	147	5	ADL72350	ADL72350 Human ova

ALIGNMENTS

RESULT 1
ABA70710/c
ID ABA70710 standard; DNA, 180 BP.
XX
AC ABA70710;
DT 01-FEB-2002 (first entry)
XX
DE Human foetal liver single exon nucleic acid probe #19015.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000669.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX PA
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483447/52.
XX
DR Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human fetal liver.
PT
PS Claim 4; SEQ ID NO 19015; 639bp + Sequence Listing; English.
XX
XX The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human foetal liver. The
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human foetal liver. The

CC present sequence is a single exon nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 180 BP; 49 A; 42 C; 25 G; 64 T; 0 U; 0 Other;

Query Match 13.9%; Score 176.8; DB 4; Length 180;
Best Local Similarity 98.9%; Pred. No. 2.4e-43;
Matches 178; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 787 CAAGGCGCACAGAACATCAGTGAAGCAATTTACCGAAGATTGGTCAGAGAGTGTCT 846
DB 180 CAAGGCGCACAGAACATCAGTGAAGCAATTTACCGAAGATTGGTCAGAGAGAGTGTCT 121

QY 847 CAACATGGCTTTGTGCGACAAGATTTAAAGATCTTGTGGATTTTCAAGATGAATACA 906
DB 120 CAACATGGCTTTGTGCGACAAGATTTAAAGATCTTGTGGATTTTCAAGATGAATACA 61

QY 907 TTGATGGAAGAAAGACTGTGAATCTTACAAAGAAAGTGTGGTGAATTTGAAAAATTG 966
DB 60 TTGATGGAAGAAAGACTGTGAATCTTACAAAGAAAGTGTGGTGAATTTGAAAAATTG 1

RESULT 2
AAI50879/c
ID AAI50879 standard; DNA; 180 BP.

XX AAI50879;

DT 17-OCT-2001 (first entry)

DE Probe #19565 used to measure gene expression in human placenta sample.

KM Probe: microarray; human; placenta; antenatal diagnosis;
KW genetic disorder; ss.

XX Homo sapiens.

PN W0200157272-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US000663.

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PA Penn SG, Hanzel DK, Chen W, Rank DR;

PS WPI; 2001-488897/53.

PT Human genome-derived single exon nucleic acid probes useful for analyzing

XX gene expression in human placenta.

PS Claim 25; SEQ ID NO 19565; 654bp; English.

XX The present invention relates to single exon nucleic acid probes (SENP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders

XX Sequence 180 BP; 49 A; 42 C; 25 G; 64 T; 0 U; 0 Other;

Query Match 13.9%; Score 176.8; DB 4; Length 180;

Best Local Similarity 98.9%; Pred. No. 2.4e-43;
Matches 178; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 787 CAAGGCGCACAGAACATCAGTGAAGCAATTTACCGAAGATTGGTCAGAGAGTGTCT 846
DB 180 CAAGGCGCACAGAACATCAGTGAAGCAATTTACCGAAGATTGGTCAGAGAGAGTGTCT 121

QY 847 CAACATGGCTTTGTGCGACAAGATTTAAAGATCTTGTGGATTTTCAAGATGAATACA 906
DB 120 CAACATGGCTTTGTGCGACAAGATTTAAAGATCTTGTGGATTTTCAAGATGAATACA 61

QY 907 TTGATGGAAGAAAGACTGTGAATCTTACAAAGAAAGTGTGGTGAATTTGAAAAATTG 966
DB 60 TTGATGGAAGAAAGACTGTGAATCTTACAAAGAAAGTGTGGTGAATTTGAAAAATTG 1

RESULT 3
AAK44908/c
ID AAK44908 standard; DNA; 180 BP.

XX AAK44908;

DT 06-NOV-2001 (first entry)

DE Human bone marrow expressed single exon probe SEQ ID NO: 19465.

KM Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukemia; lymphoma; myeloma; ss.

XX Homo sapiens.

PN W0200157276-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US000668.

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PA Penn SG, Hanzel DK, Chen W, Rank DR;

PS WPI; 2001-488900/53.

PT Human genome-derived single exon nucleic acid probes useful for analyzing

XX gene expression in human bone marrow.

PS Example 4; SEQ ID NO 19465; 658bp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid

CC probes which are derived from genomic sequences expressed in the human

CC bone marrow. They can be used to measure gene expression in bone marrow

CC samples, which may enable the improved diagnosis and treatment of cancers

CC such as lymphoma, leukemia and myeloma. The present sequence is one of
CC the probes of the invention

XX Sequence 180 BP; 49 A; 42 C; 25 G; 64 T; 0 U; 0 Other;

Query Match 13.9%; Score 176.8; DB 4; Length 180;
Best Local Similarity 98.9%; Pred. No. 2.4e-43;
Matches 178; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 787 CAAGGCGCACAGAACATCAGTGAAGCAATTTACCGAAGATTGGTCAGAGAGTGTCT 846
DB 180 CAAGGCGCACAGAACATCAGTGAAGCAATTTACCGAAGATTGGTCAGAGAGAGTGTCT 121

QY 847 CAACATGCTTTGTGCAACAAGATTAAAGACTTTGTTGATTTTCAAGATGATTAACA 906
DB 120 CAACATGCTTTGTGCAACAAGATTAAAGACTTTGTTGATTTTCAAGATGATTAACA 61
QY 907 TTGATGAAAAAGAACTGTTGAATCTTACAAAGAAAGCTGGCTGATGATTTGAAAATTG 966
DB 60 TTGATGAAAAAGAACTGTTGAATCTTACAAAGAAAGCTGGCTGATGATTTGAAAATTG 1

RESULT 4

AAK18959/c
ID AAK18959 standard; DNA; 180 BP.

AC AAK18959;

DT 05-NOV-2001 (first entry)

DE Human brain expressed single exon probe SEQ ID NO: 18950.

XX Human; brain expressed exon; gene expression analysis; probe; microarray;
KM Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
KW 68.

XX Homo sapiens.

OS MO200157275-A2.

PN 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US000667.

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483446/52.

XX Single exon nucleic acid probes for analyzing gene expression in human
PT brains.

XX Example 4; SEQ ID NO 18950; 650pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention

XX Sequence 180 BP; 49 A; 42 C; 25 G; 64 T; 0 U; 0 Other;

Query Match 13.9%; Score 176.8; DB 4; Length 180;

Best Local Similarity 98.9%; Pred. No. 2.4e-43;

Matches 178; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 787 CAAGGCGCACAGAACATCAGCTGAAGCAATTTACCGAAGATTGTCAGAGAGTGTCT 846

DB 180 CAAGGCGCACAGAACATCAGCTGAAGCAATTTACCGAAGATTGTCAGAGAGATGTCT 121

QY 847 CAACATGCTTTGTGCAACAAGATTAAAGACTTTGTTGATTTTCAAGATGATTAACA 906

DB 120 CAACATGCTTTGTGCAACAAGATTAAAGACTTTGTTGATTTTCAAGATGATTAACA 61

QY 907 TTGATGAAAAAGAACTGTTGAATCTTACAAAGAAAGCTGGCTGATGATTTGAAAATTG 966

DB 60 TTGATGAAAAAGAACTGTTGAATCTTACAAAGAAAGCTGGCTGATGATTTGAAAATTG 1

RESULT 5

AB844572/c
ID AB844572 standard; DNA; 180 BP.

AC AB844572;

DT 25-FEB-2003 (first entry)

DE Human liver single exon probe, SEQ ID NO 19562.

XX Human; single exon nucleic acid probe; liver; cirrhosis;
KM hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW coronary heart disease; ss.

XX Homo sapiens.

PN MO200157273-A2.

PF 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US000664.

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488898/53.

XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human adult liver.

XX Claim 4; SEQ ID NO 19562; 658pp; English.

XX The invention relates to a single exon nucleic acid probe (SENAP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given at high
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (I) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. AB825011-AB851005 represent human
CC liver single exon nucleic acid probes of the invention. Note: The
CC sequence information for this patent does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 180 BP; 49 A; 42 C; 25 G; 64 T; 0 U; 0 Other;

Query Match 13.9%; Score 176.8; DB 4; Length 180;

Best Local Similarity 98.9%; Pred. No. 2.4e-43;

Matches 178; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 787 CAAGGCGCACAGAACATCAGCTGAAGCAATTTACCGAAGATTGTCAGAGAGTGTCT 846

DB 180 CAAGGCGCACAGAACATCAGCTGAAGCAATTTACCGAAGATTGTCAGAGAGATGTCT 121

QY 847 CAACATGCTTTGTGCAACAAGATTAAAGACTTTGTTGATTTTCAAGATGATTAACA 906

DB 120 CAACATGCTTTGTGCAACAAGATTAAAGACTTTGTTGATTTTCAAGATGATTAACA 61

QY 907 TTGATGAAAAGAACTGTTGAATCTTACAAAAGAGTCTGCGTGAATTTGAAAATTG 966
DB 60 TTGATGAAAAGAACTGTTGAATCTTACAAAAGAGTCTGCGTGAATTTGAAAATTG 1

RESULT 6
ABSI19150/c
ID ABSI19150 standard; DNA; 180 BP.
XX
AC ABSI19150;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human genome-derived single exon probe ORF from lung SEQ ID NO 19141.
XX
KW Human; db; single exon probe; asthma; lung cancer; COPD; ILD;
KW Chronic obstructive pulmonary disease; interstitial lung disease;
KW Familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberosus sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease; open reading frame; ORF.
XX
OS Homo sapiens.
XX
PN WO200186003-A2.
XX
PD 15-NOV-2001.
XX
PF 30-JAN-2001; 2001WO-US000665.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2002-114183/15.
XX
PT Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples.
XX
PS Claim 4; SEQ ID NO 19141; 634bp; English.
XX
XX The invention relates to a spatially-addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human lung comprising single exon nucleic acid probes having one of
XX 12614 nucleic acid sequences mentioned in the specification, or their
XX complements or the 12387 open reading frames derived from the 12614
XX probes. Also included are a microarray comprising the novel set of probes
XX; the novel set of probes which hybridize at high stringency to a nucleic
XX acid expressed in the human lung; measuring gene expression in a sample
XX derived from human lung, comprising (a) contacting the array with a
XX collection of detectably labeled nucleic acids derived from human lung
XX mRNA, and (b) measuring the label detectably bound to each probe of the
XX array; identifying exons in a eukaryotic genome, comprising (a)
XX algorithmically predicting at least one exon from genomic sequences of
XX the eukaryote; and (b) detecting specific hybridization of detectably
XX labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
XX having a fragment identical to the predicted exon, the probe is included
XX in the above mentioned microarray; assigning exons to a single gene,
XX comprising (a) identifying exons from genomic sequence by the method
XX above and (b) measuring the expression of each of the exons in several
XX tissues and/or cell types using hybridization to a single exon

CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene expression
CC analysis, and for identifying exons in a gene, particularly using human
CC lung derived mRNA and for the study of lung diseases such as asthma, lung
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC tuberosus sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a single exon probe open reading frame of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WPI at ftp.wpi.int/pub/published_pcc_sequences
XX
SQ Sequence 180 BP; 49 A; 42 C; 25 G; 64 T; 0 U; 0 Other;
XX
Query Match 13.9%; Score 176.8; DB 6; Length 180;
Best Local Similarity 98.9%; Pred. No. 2,4e-43;
Matches 178; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 787 CAAGCGCACAGAAATCATGAGTGAAGCAATTTACCGAAGATTGTCAGAGAGTGTCT 846
DB 180 CAAGCGCACAGAAATCATGAGTGAAGCAATTTACCGAAGATTGTCAGAGAGAGTGTCT 121
QY 847 CAACATGCGCTTTGTGCAACAAGATTAAAGATCTTGTGTATTTTCAAGATGAATGA 906
DB 120 CAACATGCGCTTTGTGCAACAAGATTAAAGATCTTGTGTATTTTCAAGATGAATGA 61
QY 907 TTGATGAAAAGAACTGTTGAATCTTACAAAAGAGTCTGCGTGAATTTGAAAATTG 966
DB 60 TTGATGAAAAGAACTGTTGAATCTTACAAAAGAGTCTGCGTGAATTTGAAAATTG 1

RESULT 7
AAI13283/c
ID AAI13283 standard; DNA; 132 BP.
XX
XX AAI13283;
XX
AC AAI13283;
XX
DT 12-OCT-2001 (first entry)
XX
DE Probe #13216 for gene expression analysis in human cervical cell sample.
XX
KW Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer; ss.
XX
OS Homo sapiens.
XX
XX WO200157278-A2.
XX
PN 09-AUG-2001.
XX
PD 30-JAN-2001; 2001WO-US000670.
XX
PF 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488901/53.
XX

PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human cervical epithelial cells.
XX
PS Claim 25; SEQ ID NO 13216; 487bp; English.
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SNP). The present sequence is one such probe. The SNPs are derived
CC from human HeLa cells. The SNPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging of
CC diseases of the cervix, notably cervical cancer. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 132 BP; 35 A; 29 C; 24 G; 44 T; 0 U; 0 Other;
Query Match 10.4%; Score 132; DB 4; Length 132;
Best Local Similarity 100.0%; Pred. No. 1.1e-29;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 976 GACTGCGTAGTAAAGTCTGAGAAAATTGAAAGCTCAGACCAAGTTAAATCCCTTT 1035
DB 132 GACTGCGTAGTAAAGTCTGAGAAAATTGAAAGCTCAGACCAAGTTAAATCCCTTT 73
QY 1036 CTTGAGGAATTCCTGATGAATTTATATGTCCAACTAAGAGAACTTATGAAGATCCGG 1095
DB 72 CTTGAGGAATTCCTGATGAATTTATATGTCCAACTAAGAGAACTTATGAAGATCCGG 13
QY 1096 TCATCGCATCAG 1107
DB 12 TCATCGCATCAG 1
RESULT 8
ABA68387/c
ID ABA68387 standard; DNA; 132 BP.
XX
AC ABA68387;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human foetal liver single exon nucleic acid probe #16692.
XX
KM Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000669.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
PT WPI; 2001-483447/52.
XX
DR WPI; 2001-483447/52.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human fetal liver.
PS Claim 4; SEQ ID NO 16692; 639bp + Sequence Listing; English.

XX
CC The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human foetal liver. The
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human fetal liver. The
CC present sequence is a single exon nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 132 BP; 35 A; 29 C; 24 G; 44 T; 0 U; 0 Other;
Query Match 10.4%; Score 132; DB 4; Length 132;
Best Local Similarity 100.0%; Pred. No. 1.1e-29;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 976 GACTGCGTAGTAAAGTCTGAGAAAATTGAAAGCTCAGACCAAGTTAAATCCCTTT 1035
DB 132 GACTGCGTAGTAAAGTCTGAGAAAATTGAAAGCTCAGACCAAGTTAAATCCCTTT 73
QY 1036 CTTGAGGAATTCCTGATGAATTTATATGTCCAACTAAGAGAACTTATGAAGATCCGG 1095
DB 72 CTTGAGGAATTCCTGATGAATTTATATGTCCAACTAAGAGAACTTATGAAGATCCGG 13
QY 1096 TCATCGCATCAG 1107
DB 12 TCATCGCATCAG 1
RESULT 9
AA148602/c
ID AA148602 standard; DNA; 132 BP.
XX
AC AA148602;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #17288 used to measure gene expression in human placenta sample.
XX
KM Probe; microarray; human; placenta; antenatal diagnosis;
XX
KM genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000663.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
PT WPI; 2001-488897/53.
XX
DR WPI; 2001-488897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human placenta.
PS Claim 25; SEQ ID NO 17288; 654bp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SNP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful

CC for antenatal diagnosis of human genetic disorders
XX Sequence 132 BP; 35 A; 29 C; 24 G; 44 T; 0 U; 0 Other;
SQ Sequence 132 BP; 35 A; 29 C; 24 G; 44 T; 0 U; 0 Other;
Query Match 10.4%; Score 132; DB 4; Length 132;
Best Local Similarity 100.0%; Pred. No. 1.1e-29;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 976 GACTGCGTAGTAAAGTGTGAGGAAATTGAGAGCTCAGACCAAGTTAAATCCCTT 1035
DB 132 GACTGCGTAGTAAAGTGTGAGGAAATTGAGAGCTCAGACCAAGTTAAATCCCTT 73
QY 1036 CTTGAGGAATTCCTGATGAATTATATGTCCAATACTAGAGAATTATGAAAGATCCG 1095
DB 72 CTTGAGGAATTCCTGATGAATTATATGTCCAATACTAGAGAATTATGAAAGATCCG 13
QY 1096 TCATCGCATCAG 1107
DB 12 TCATCGCATCAG 1
RESULT 10
ABA50437/C
ID ABA50437 standard; DNA; 132 BP.
XX ABA50437;
XX 01-FEB-2002 (first entry)
XX DT
XX DE Human breast cell single exon nucleic acid probe #9132.
XX KM Human; microarray; single exon probe; gene expression; breast; disease;
XX KM cancer; ss.
XX OS Homo sapiens.
XX PN WO200157271-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000662.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX PT WPI; 2001-496933/54.
XX PT New spatially-addressable set of single exon nucleic acid probes, useful
XX PT for measuring gene expression in sample derived from human breast,
XX PT comprises number of single exon nucleic acid probes.
XX PS Claim 4; SEQ ID NO 9132; 327bp + Sequence Listing; English.
XX XX The invention relates to a spatially-addressable set of single exon
XX CC nucleic acid probes for measuring gene expression in a sample derived
XX CC from human breast and BT 474 cells. The method involves contacting the
XX CC probes with a collection of detectably labelled nucleic acids derived
XX CC from mRNA of human breast, and then measuring the label bound to each
XX CC probe of the microarray. The probes are useful for verifying the
XX CC expression of regions of genomic DNA predicted to encode proteins. They
XX CC are useful for gene discovery, and for determining predisposition and/or
XX CC prognosing breast disease. Gene expression analysis is useful for
XX CC assessing the toxicity of chemical agents on cells. The microarray of
XX CC this invention presents a far greater diversity of probes for measuring

CC gene expression, with far less bias than expressed sequence tag
CC microarrays. The method is suitable for rapid production of functional
CC information from genomic sequence. The present sequence is a single exon
CC nucleic acid probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp://ipo.int/pub/published_pct_sequences
XX SQ Sequence 132 BP; 35 A; 29 C; 24 G; 44 T; 0 U; 0 Other;
Query Match 10.4%; Score 132; DB 4; Length 132;
Best Local Similarity 100.0%; Pred. No. 1.1e-29;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 976 GACTGCGTAGTAAAGTGTGAGGAAATTGAGAGCTCAGACCAAGTTAAATCCCTT 1035
DB 132 GACTGCGTAGTAAAGTGTGAGGAAATTGAGAGCTCAGACCAAGTTAAATCCCTT 73
QY 1036 CTTGAGGAATTCCTGATGAATTATATGTCCAATACTAGAGAATTATGAAAGATCCG 1095
DB 72 CTTGAGGAATTCCTGATGAATTATATGTCCAATACTAGAGAATTATGAAAGATCCG 13
QY 1096 TCATCGCATCAG 1107
DB 12 TCATCGCATCAG 1
RESULT 11
AAK42535/C
ID AAK42535 standard; DNA; 132 BP.
XX AAK42535;
XX 06-NOV-2001 (first entry)
XX DT
XX DE Human bone marrow expressed single exon probe SEQ ID NO: 17092.
XX KM Human; bone marrow expressed exon; gene expression analysis; probe;
XX KM microarray; cancer; leukemia; lymphoma; myeloma; ss.
XX OS Homo sapiens.
XX PN WO200157276-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000668.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX PT WPI; 2001-488900/53.
XX PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX PT gene expression in human bone marrow.
XX PS Example 4; SEQ ID NO 17092; 658bp + Sequence Listing; English.
XX XX The present invention provides a number of single exon nucleic acid
XX CC probes which are derived from genomic sequences expressed in the human
XX CC bone marrow. They can be used to measure gene expression in bone marrow
XX CC samples, which may enable the improved diagnosis and treatment of cancers
XX CC such as lymphoma, leukemia and myeloma. The present sequence is one of
XX CC the probes of the invention

XX Sequence 132 BP; 35 A; 29 C; 24 G; 44 T; 0 U; 0 Other;
SQ
Query Match 10.4%; Score 132; DB 4; Length 132;
Best Local Similarity 100.0%; Pred.No.1.1e-29;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 976 GACTGCGTAGTAAAGTCTGAGGAAATTTGAAGAGCTCAGGACCAAGTTAAATCCCTTT 1035
DB 132 GACTGCGTAGTAAAGTCTGAGGAAATTTGAAGAGCTCAGGACCAAGTTAAATCCCTTT 73
QY 1036 CTTGAGGAATTCCTGATGAATTTATATGTCCAATTAAGAACTTATGAAAGATCCGG 1095
DB 72 CTTGAGGAATTCCTGATGAATTTATATGTCCAATTAAGAACTTATGAAAGATCCGG 13
QY 1096 TCATCGCATCAG 1107
DB 12 TCATCGCATCAG 1
RESULT 12
ABS42143/C
ID ABS42143 standard; DNA; 132 BP.
XX
AC ABS42143;
XX
DT 25-FEB-2003 (first entry)
XX
DE Human liver single exon probe. SEQ ID NO 17133.
XX
KM Human; single exon nucleic acid probe; liver; cirrhosis;
KM hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KM coronary heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157273-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000664.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488896/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human adult liver.
XX
PS Claim 4; SEQ ID NO 17133; 658bp; English.
XX
PS The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (It may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic, liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABS25011-ABS51005 represent human
CC liver single exon nucleic acid probes of the invention. Note: The
CC sequence information for this patent does not appear in the printed

CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 132 BP; 35 A; 29 C; 24 G; 44 T; 0 U; 0 Other;
Query Match 10.4%; Score 132; DB 4; Length 132;
Best Local Similarity 100.0%; Pred.No.1.1e-29;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 976 GACTGCGTAGTAAAGTCTGAGGAAATTTGAAGAGCTCAGGACCAAGTTAAATCCCTTT 1035
DB 132 GACTGCGTAGTAAAGTCTGAGGAAATTTGAAGAGCTCAGGACCAAGTTAAATCCCTTT 73
QY 1036 CTTGAGGAATTCCTGATGAATTTATATGTCCAATTAAGAACTTATGAAAGATCCGG 1095
DB 72 CTTGAGGAATTCCTGATGAATTTATATGTCCAATTAAGAACTTATGAAAGATCCGG 13
QY 1096 TCATCGCATCAG 1107
DB 12 TCATCGCATCAG 1
RESULT 13
AA108924/C
ID AA108924 standard; DNA; 132 BP.
XX
AC AA108924;
XX
DT 09-OCT-2001 (first entry)
XX
DE Probe #8915 used to measure gene expression in human breast sample.
XX
KM Probe; human; breast disease; breast cancer; development disorder; ss;
KM inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS Homo sapiens.
XX
PN WO200157270-A2.
XX
PD 09-AUG-2001.
XX
PF 29-JAN-2001; 2001WO-US000661.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-476286/51.
XX
PT Novel single exon nucleic acid probe used to measuring gene expression in
PT a human breast.
XX
PS Claim 25; SEQ ID NO 8915; 322bp; English.
XX
PS The present invention relates to novel single exon nucleic acid probes.
CC The present sequence is one such probe. The probes are useful for
CC measuring human gene expression in a human breast sample, where the probe
CC hybridises at high stringency to a nucleic acid expressed in the human
CC breast. The probes are useful for predicting, diagnosing, grading,
CC staging, monitoring and prognosing diseases of the human breast,
CC particularly those diseases with polygenic aetiology. The diseases
CC include: breast cancer, disorders of development, inflammatory diseases
CC of the breast, fibrocystic changes, proliferative breast disease and non-
CC carcinoma tumours. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format

CC directly from WIP0 at ftp.wipo.int/pub/published_pct_sequences
XX Sequence 132 BP; 35 A; 29 C; 24 G; 44 T; 0 U; 0 Other;
SQ
Query Match 10.4%; Score 132; DB 5; Length 132;
Best Local Similarity 100.0%; Pred. No. 1.1e-29;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 976 GACTGCGTAGTAAAGTCTGAGGAAATTTGAAGAGCTCAGAGCCAGGTTAAATCCCTTT 1035
DB 132 GACTGCGTAGTAAAGTCTGAGGAAATTTGAAGAGCTCAGAGCCAGGTTAAATCCCTTT 73
QY 1036 CTTGAGGAATTCCTGATGAATTTATATGTCCTCAATTAAGAACTTATGAAGATCCGG 1095
DB 72 CTTGAGGAATTCCTGATGAATTTATATGTCCTCAATTAAGAACTTATGAAGATCCGG 13
QY 1096 TCATCGCATCAG 1107
DB 12 TCATCGCATCAG 1
RESULT 14
ABS16579/C
ID ABS16579 standard; DNA; 132 BP.
XX
AC ABS16579;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human genome-derived single exon probe ORF from lung SEQ ID NO 16570.
XX
KW Human; db; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karsagenet syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease; open reading frame; ORF.
XX
OS Homo sapiens.
XX
PN WO200186003-A2.
XX
PD 15-NOV-2001.
XX
PF 30-JAN-2001; 2001WO-US000665.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2002-114183/15.
XX
PT Spatially-addressable set of single exon nucleic acid probes, used to
XX measure gene expression in human lung samples.
XX
PS Claim 4; SEQ ID NO 16570; 634bp; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614

CC probes. Also included are a microarray comprising the novel set of probes
CC; the novel set of probes which hybridise at high stringency to a nucleic
CC acid expressed in the human lung; measuring gene expression in a sample
CC derived from human lung, comprising (a) contacting the array with a
CC collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of the
CC array; identifying exons in a eukaryotic genome, comprising (a)
CC algorithmically predicting at least one exon from genomic sequences of
CC the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene expression
CC analysis, and for identifying exons in a gene, particularly using human
CC lung derived mRNA and for the study of lung diseases such as asthma, lung
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
CC Karsagenet syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a single exon probe open reading frame of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIP0 at ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 132 BP; 35 A; 29 C; 24 G; 44 T; 0 U; 0 Other;
Query Match 10.4%; Score 132; DB 6; Length 132;
Best Local Similarity 100.0%; Pred. No. 1.1e-29;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 976 GACTGCGTAGTAAAGTCTGAGGAAATTTGAAGAGCTCAGAGCCAGGTTAAATCCCTTT 1035
DB 132 GACTGCGTAGTAAAGTCTGAGGAAATTTGAAGAGCTCAGAGCCAGGTTAAATCCCTTT 73
QY 1036 CTTGAGGAATTCCTGATGAATTTATATGTCCTCAATTAAGAACTTATGAAGATCCGG 1095
DB 72 CTTGAGGAATTCCTGATGAATTTATATGTCCTCAATTAAGAACTTATGAAGATCCGG 13
QY 1096 TCATCGCATCAG 1107
DB 12 TCATCGCATCAG 1
RESULT 15
AAH36289
ID AAH36289 standard; cDNA; 92 BP.
XX
AC AAH36289;
XX
DT 03-SEP-2001 (first entry)
XX
DE Human colon cancer antigen encoding cDNA SEQ ID NO:3371.
XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma; 88.
XX
OS Homo sapiens.
XX
PN WO200122920-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-US026524.

XX 29-SEP-1999; 99US-0157137P.
PR 03-NOV-1999; 99US-0163280P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Barash SC, Birse CE, Rosen CA;
XX WPI; 2001-235357/24.
DR P-PSDB; AAG76884.
XX
PT Nucleic acid encoding 4277 human colon cancer-associated polypeptides,
XX useful for preventing, diagnosing and/or treating colorectal cancers.
XX
PS Claim 1; Page 5177-5178; 9803pp; English.
XX
CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where the
CC proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene therapy
CC and vaccine production. N and P may be used in the prevention, diagnosis
CC and treatment of diseases associated with inappropriate P expression. For
CC example, N and P may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of P by expressing inactive proteins or to
CC supplement the patient's own production of P. Additionally, N may be used
CC to produce the colon cancer-associated Ps, by inserting the nucleic acids
CC into a host cell and culturing the cell to express the proteins. N and P
CC can be used in the prevention, diagnosis and treatment of colorectal
CC carcinomas and cancers. AAH37196 to AAH37204 and AAB77789 represent
CC sequences used in the exemplification of the present invention. N.B.
CC Pages 666 to 682 and page 7053 of the sequence listing were missing at
CC time of publication, meaning no sequences are present for SEQ ID NO:1027
CC to 1052, 7921 and 7922
XX
SQ Sequence 92 BP; 27 A; 9 C; 23 G; 28 T; 0 U; 5 Other;

Query Match 6.8%; Score 87; DB 4; Length 92;
Best Local Similarity 94.6%; Pred. No. 5.3e-16;
Matches 87; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 591 GGTGATTTAAACAGTGTGGAGTGAATAAATGAGTGTCTGCACTAGTGAATAAAGCACATGAT 650
DB 1 GGTGATTTAAACAGTGTGGAGTGAATAAATGAGTGTCTGCACTAGTGAATAAAGCACATGAT 60
QY 651 CTTGGAATTAACCTGCTGCGATTTCTTCTCACA 682
DB 61 CTTGGAATTAACCTGCTGCGATTTCTTCTTNNANA 92

Search completed: February 5, 2005, 08:47:55
Job time : 760 secs

This Page Blank (uspto)

QY 72 CCCCCCTCTGAGGCTGTTTTTTCTTCAATAAAGAACTGGTAACTGATTCACACA 131
 DB 129 CCCCCCTCTGAGGCTGTTTTTTCTTCAATAAAGAACTGGTAACTGATTCACACA 188
 QY 132 TTAGCTGATCAGTGAACGATGTCACACTGCTGCTCTTCCCTTTCCCTTGGGCTACT 191
 DB 189 TTAGCTGATCAGTGAACGATGTCACACTGCTGCTCTTCCCTTTCCCTTGGGCTACT 248
 QY 192 TGTCTCTTGGACAAAACAATTCCCTGTATCTCGTTACGTACTTACTGAAGCCACAT 251
 DB 249 TGTCTCTTGGACAAAACAATTCCCTGTATCTCGTTACGTACTTACTGAAGCCACAT 308
 QY 252 TCTCCATTGAAGTTTCATACCTATGCTGTCACCTGCTGCTGTTTCTCCCTTCAAGACAT 311
 DB 309 TCTCCATTGAAGTTTCATACCTATGCTGTCACCTGCTGCTGTTTCTCCCTTCAAGACAT 368
 QY 312 ATTTTGGCATCGTGTTCACAGATGTCACACTGCTCTTATGAAATCTGAAAATGACAG 371
 DB 369 ATTTTGGCATCGTGTTCACAGATGTCACACTGCTCTTATGAAATCTGAAAATGACAG 428
 QY 372 ATGCTGACGATGATGGAACAGCCTAAGTGGAGCCCTGTGAGGGTTTGGCCAGTTTCCCA 431
 DB 429 ATGCTGACGATGATGGAACAGCCTAAGTGGAGCCCTGTGAGGGTTTGGCCAGTTTCCCA 488
 QY 432 GACTCCACGTTTGGCATCAGGGGAGCTGATGGAAGTGGTGGTTTGGTGAATGACAG 491
 DB 489 GACTCCACGTTTGGCATCAGGGGAGCTGATGGAAGTGGTGGTTTGGTGAATGACAG 548
 QY 492 TCATCAAAATTATATAGATGTGTAAGTGTAAAGATGCTCTTGGCGGATGTCATTT 551
 DB 549 TCATCAAAATTATATAGATGTGTAAGTGTAAAGATGCTCTTGGCGGATGTCATTT 608
 QY 552 TCCTCAATGAGAGCTCTTGTCTCACTGCTCCCTCATGAGTATTAACAGTGGGAT 611
 DB 609 TCCTCAATGAGAGCTCTTGTCTCACTGCTCCCTCATGAGTATTAACAGTGGGAT 668
 QY 612 GATMAAATGAGTGTCTGATGTAAGAAAGACATGATCTTGAATTAACCTGCTGGAT 671
 DB 669 GATMAAATGAGTGTCTGATGTAAGAAAGACATGATCTTGAATTAACCTGCTGGAT 728
 QY 672 TTTTCTTCAAGCAGCTTTGTATGAGAAACAAGTCTTCAATTTTTCAGCTGGCATCA 731
 DB 729 TTTTCTTCAAGCAGCTTTGTATGAGAAACAAGTCTTCAATTTTTCAGCTGGCATCA 788
 QY 732 TGTGTCAGAGTGGCCAGTCAAAATTTGGATTGTTCTTTTCCCATATCTT----- 784
 DB 789 TGTGTCAGAGTGGCCAGTCAAAATTTGGATTGTTCTTTTCCCATATCTTATGCTTTT 848
 QY 785 ----- 784
 DB 849 GAATTAATAATAAAGTACACTGAGTGGGACCTGCTCTGTTGCTTGTGCTTTT 908
 QY 785 ----- 784
 DB 909 TCCCATGATGGGAGATGCTAGTCTCAGGGTCTAGTGAATAGTCTGATATATATGAT 968
 QY 785 ----- 784
 DB 969 ACTAATACTGAGAAATATCTTCAACATTTGACTCAGACACAGGATATGTCACAACCTGT 1028
 QY 785 ----- 784
 DB 1029 GCTTTGACCTAATACCTTTTACTTGTACTGTTTCAATGACAAAACAGTGAACATC 1088
 QY 785 ----- AGCAAGGCGCACAGAACATCAGCTGAAGCAA 815
 DB 1089 TGGCAATTGACCTGGAAACATTTTGGCAAGAGGCGCACAGAACATCAGCTGAAGCAA 1148
 QY 816 TTATACGAAGATTGCTCAGAGAGTGTCTCAACATGCTTTGTGTCAGAAATTTTAAA 875
 DB 1149 TTATACGAAGATTGCTCAGAGAGTGTCTCAACATGCTTTGTGTCAGAAATTTTAAA 1208

QY 876 GATCTTGTGGTATTTTCAAGATGAATTAACATTGATGAGAAAAGAACTGTTGAATCTTACA 935
 DB 1209 GATCTTGTGGTATTTTCAAGATGAATTAACATTGATGAGAAAAGAACTGTTGAATCTTACA 1268
 QY 936 AAGAAGCTGGCTGATGATTTGAATTTGAATCTCTGAGACTGGCTGTAAAGCTG 995
 DB 1269 AAGAAGCTGGCTGATGATTTGAATTTGAATCTCTGAGACTGGCTGTAAAGCTG 1328
 QY 996 AGGAAATTTGAAGAGCTCAGACCAAGTTAAATCCCTTCTTCAAGAAATTCCTGATGAA 1055
 DB 1329 AGGAAATTTGAAGAGCTCAGACCAAGTTAAATCCCTTCTTCAAGAAATTCCTGATGAA 1388
 QY 1056 TTTATATGTCATAATCTAGAGAACTTATGAAGATCCGTCATCGCATCAGATGCTAT 1115
 DB 1389 TTTATATGTCATAATCTAGAGAACTTATGAAGATCCGTCATCGCATCAGATGCTAT 1448
 QY 1116 TCATATGAAAAGAAAGCAATGCAAAATTTGGATCAGAAAAAGAAAGCTCAAGTCCCATG 1175
 DB 1449 TCATATGAAAAGAAAGCAATGCAAAATTTGGATCAGAAAAAGAAAGCTCAAGTCCCATG 1508
 QY 1176 ACAATCTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1235
 DB 1509 ACAATCTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1568
 QY 1236 AATGATGCTGGAGACACACCAAAAGTAAA 1266
 DB 1569 AATGATGCTGGAGACACACCAAAAGTAAA 1599

RESULT 2
 LOCUS CR619104 1600 bp mRNA linear HTC 21-JUL-2004
 DEFINITION full-length cDNA clone CS01075YB08 of Placenta Cot 25-normalized
 of Homo sapiens (human).
 ACCESSION CR619104
 VERSION CR619104.1 GI:50499911
 KEYWORDS HTC; CNSLT_cDNA.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 1 (bases 1 to 1600)
 Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 Full-length cDNA libraries and normalization.
 JOURNAL Unpublished
 REMARK Contact : Feng Liang Email : fliang@life.rockefeller.edu
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue
 2 (bases 1 to 1600)

REFERENCE
 Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)

COMMENT
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with NotI and cloned
 into the NotI and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen.
 FEATURES
 Location/Qualifiers
 1..1600

source
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS01075YB08"
 /issue_type="Placenta Cot 25-normalized"
 /plasmid="pCMVSPORT_6"

ORIGIN

Query Match 61.5%; Score 782; DB 3; Length 1600;
 Best Local Similarity 77.3%; Pred. No. 1,6e-212; Indels 135; Gaps 5;
 Matches 1075; Conservative 0; Mismatches 180;


```

Qy 12 TCACCTGCGGCGACGTGACCCGCGCCGCGGACCTTGAAGCGGATCCCGCGC 71
Db 13 TCACCTGCGGCGACGTGACCCGCGCCGCGGACCTTGAAGCGGATCCCGCGC 72
Qy 72 CCCCCGCTCTGACAGGCTGTTTTTCTTCAATAAAGAAACATGTGAACTGATTCACACA 131
Db 73 CCCCCGCTCTGACAGGCTGTTTTTCTTCAATAAAGAAACATGTGAACTGATTCACACA 132
Qy 132 TTAGCTGATGATGCTGAGAGATGCACTGCTGTCCTCTTCCCTTCCCTTGGCTACT 191
Db 133 TTAGCTGATGATGCTGAGAGATGCACTGCTGTCCTCTTCCCTTCCCTTGGCTACT 192
Qy 192 TGCTCTTGACAAACAAATTCGCTGACTGCTTACGTCGATGATGATGATGATGATGAT 251
Db 193 TGCTCTTGACAAACAAATTCGCTGACTGCTTACGTCGATGATGATGATGATGATGAT 252
Qy 252 TCTTCATTTGAAGTTTCATACCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 311
Db 253 TCTTCATTTGAAGTTTCATACCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 312
Qy 312 ATTTTGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 371
Db 313 ATTTTGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 372
Qy 372 ATGCTGCGAGTGTGACAAACAGCTAGTGGCAGCCCTGAGAGGTTTGGCAGTTTCCCA 431
Db 373 ATGCTGCGAGTGTGACAAACAGCTAGTGGCAGCCCTGAGAGGTTTGGCAGTTTCCCA 432
Qy 432 GACTCCACGCTGTTGGCATCAGGGGAGCTGATGAACTGTGTTTGGTGGATGACAG 491
Db 433 GACTCCACGCTGTTGGCATCAGGGGAGCTGATGAACTGTGTTTGGTGGATGACAG 492
Qy 492 TCATACAAATATATAGATGTGTAGTGTAAAGTGGCTCTTGGCGGACATGTCATTT 551
Db 493 TCATACAAATATATAGATGTGTAGTGTAAAGTGGCTCTTGGCGGACATGTCATTT 552
Qy 552 TCTCCTATGAGACCTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 611
Db 553 TCTCCTATGAGACCTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 612
Qy 612 GATTAATGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 671
Db 613 GATTAATGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 672
Qy 672 TTTTCTTACAGCCAGTTTCTGATGAGAGAAACAGTCTTCACTTTCAGTGGCATCA 731
Db 673 TTTTCTTACAGCCAGTTTCTGATGAGAGAAACAGTCTTCACTTTCAGTGGCATCA 732
Qy 732 TGTGTGAGATGAGTGCAGATCAAAATTTGATGTTTCTTACCAATCTTGAAGCAAG 791
Db 733 TGTGTGAGATGAGTGCAGATCAAAATTTGATGTTTCTTACCAATCTTGAAGCAAG 792
Qy 792 CGCAGAGATGAGTGCAGTGAAGC-----AATTT 818
Db 793 GAATTAATAATTAAGTACAGTGTGAGCACTGTCTCTGTTTGGCTTGTGCTTTT 852
Qy 819 ACCGAGATTTGTGAGAGAGGTC----- 842
Db 853 TCCCATGATGAGAGATGCTAGTCTCAAGGTCAGTGTATAGTCTGTCATATATATGAT 912
Qy 843 -----GTCCTCAACATGG 854
Db 913 ACTAATACTGAAATATATCTTACACATTTGACACACACAGGATATGACAACTTGT 972
Qy 855 CTTTGTGACAAAG-----TTTAAAGATCTTGTGTATTTTCAAGATGAATCAAT 908
Db 973 GCTTTTGACCTAATACCTCTTTTACTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1032
Qy 909 GATGAGAAAGAACTGTGATCTTACAAAGAAAGTCTGCTGATGATTTGAAATTTGAA 968
Db 1033 TGGCAATTTGACCTGAGAAACCTTTGCGACAAAGGCGCACAGAAACATCACTGAAGCA 1092
Qy 969 TCT--CTAGAGACTGCTAGTAAAGTCTGAGAGAAATTTGAAGAGCTCAGACCAAGTTA 1026

```

```

Db 1093 TTTACCGAATGTTGTCAGAGAGATGTCTCAACATGCTTTGTGCAACAAGATTTTAAA 1152
Qy 1027 AATCCCTTTCTTGAAGAAATTCCTGATGATATTTATGTCCAAATACATAGAACTTATGA 1086
Db 1153 GATCTTGTGTGTTTTCAGAGATGAAATMACATTTGAGAAAGAAAGCTTGAATCTTACA 1212
Qy 1087 A-----GATCCGCTGATCGCATGAGTGGCTATTCATATGAAAGAAAGCAATG 1136
Db 1213 AAGAAAGTCTGGCTGATGATTTGAAATTTGATGCTGCTATTCATATGAAAGAAAGCAATG 1272
Qy 1137 GAAATTTGATCAGCAAAAGAAAGAAAGTCAAGTCCCATGCAAAATCTTGTCTTCTTCA 1196
Db 1273 GAAATTTGATCAGCAAAAGAAAGAAAGTCAAGTCCCATGCAAAATCTTGTCTTCTTCA 1332
Qy 1197 GCGGTACTTACCACTAATAGACTCTGAAATATGGCCATCAATAGATGCTGAGACAC 1256
Db 1333 GCGGTACTTACCACTAATAGACTCTGAAATATGGCCATCAATAGATGCTGAGACAC 1392
Qy 1257 CAAAAGTAAA 1266
Db 1393 CAAAAGTAAA 1402

```

```

RESULT 3
AL533510 941 bp mRNA linear EST 24-MAR-2004
LOCUS AL533510 Homo sapiens ADULT BRAIN Homo sapiens cDNA clone
DEFINITION CS0DN004YJ15 5-PRIME, mRNA sequence.
ACCESSION AL533510
VERSION AL533510.3 GI:45708442
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 941)
REFERENCE 1
AUTHORS Li, W. B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 13, 2001 this sequence version replaced gi:31260591.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 Evry cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cne.fr
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 1653.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?cs=CS0DN004CE08Q1&c=1653.r.
Location/Qualifiers
1. 941

```

```

FEATURES
source
1. 941

```

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DN004YJ15"
/issue_type="ADULT BRAIN"
/dev_stage="adult"
/clone_lib="Homo sapiens ADULT BRAIN"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo (dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

```

ORIGIN

```

Query Match 60.3%; Score 767.4; DB 1; Length 941;
Best Local Similarity 98.7%; Pred. No. 2e-208;
Matches 765; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

```

QY 12 TCACCTGCGGCGACGTCAGCCGACCGCCGCGGACCTTGAAGCGGATCCCGCG 71
Db 68 TCCCTGCGCGGACGTCAGCCGACCGCCGCGGACCTTGAAGCGGATCCCGCG 127
QY 72 CCCCCCTCTGAGGCGTCTTTCTTCAATTAAGAACATGGTGAACGATTCACACA 131
Db 128 CCCCCCTCTGAGGCGTCTTTCTTCAATTAAGAACATGGTGAACGATTCACACA 187
QY 132 TTAGCTGATCAGTGAAGATGTCACACTGCTGTCCTTCTCTTCCCTCTTGACT 191
Db 188 TTAGCTGATCAGTGAAGATGTCACACTGCTGTCCTTCTCTTCCCTCTTGACT 247
QY 192 TGTCTCTTGGACAAACAAATTCGCTGTACTCGTTACGTAATTAAGTGAACG 251
Db 248 TGTCTCTTGGACAAACAAATTCGCTGTACTCGTTACGTAATTAAGTGAACG 307
QY 252 TCTCCATTGAAGTTTCACTAATGCTGTCACACTGCTGTCCTTCTCTTCCCT 311
Db 308 TCTCCATTGAAGTTTCACTAATGCTGTCACACTGCTGTCCTTCTCTTCCCT 367
QY 312 ATTTGGCATGCTGTCACAGATGTCACACTGCTGTCCTTCTCTTCCCT 371
Db 368 ATTTGGCATGCTGTCACAGATGTCACACTGCTGTCCTTCTCTTCCCT 427
QY 372 ATGCTGCGAGTGAAGACAGCTGTCAGCCCTGTCAGGCTTCCCACTTTCCCA 431
Db 428 ATGCTGCGAGTGAAGACAGCTGTCAGCCCTGTCAGGCTTCCCACTTTCCCA 487
QY 432 GACTCCACGCTGTTGGCATGAGGGGACGTCAGTGAACGTCGTTTGAAGTGA 491
Db 488 GACTCCACGCTGTTGGCATGAGGGGACGTCAGTGAACGTCGTTTGAAGTGA 547
QY 492 TCATCAAAATTAATAGATGTCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 551
Db 548 GCATCAAAATTAATAGATGTCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 607
QY 552 TCTCTCAATGAAGCTTCTTGTCTCACTGCTCTCTCATGTCGTCATTAAGTGA 611
Db 608 TCTCTCAATGAAGCTTCTTGTCTCACTGCTCTCTCATGTCGTCATTAAGTGA 667
QY 612 GATTAATGAAGTGTCTGTCATGTCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 671
Db 668 GATTAATGAAGTGTCTGTCATGTCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 727
QY 672 TTTTCTTCAAGCAGCTTCTGATGTCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 731
Db 728 TTTTCTTCAAGCAGCTTCTGATGTCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 787
QY 732 TGTGTCAGGATGTCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 786
Db 788 TGTGTCAGGATGTCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 842

RESULT 4
AL553333 889 bp mRNA linear EST 30-MAR-2004
LOCUS AL553333 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CSOD1075YB08 5-PRIME, mRNA sequence.
AL553333
ACCESSION AL553333.3 GI:45858102
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 889)
AUTHORS Li, W.-B., Gruber, C., Jeesee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 15, 2001 this sequence version replaced gi:31275147.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 Evry cedex - FRANCE

Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
1653.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CSOD1075YB08&p=1653.r.
FEATURES
source
1..889
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOD1075YB08"
/issue_type="PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 59.5%; Score 756.6; DB 1; Length 889;
Best Local Similarity 99.0%; Pred. No. 2,5e-205;
Matches 767; Conservative 4; Mismatches 3; Indels 1; Gaps 1;
QY 12 TCACCTGCGGCGACGTCAGCCGACCGCCGCGGACCTTGAAGCGGATCCCGCG 71
Db 13 TCASCTGSGCGGACGTCAGCCGACCGCCGCGGACCTTGAAGCGGATCCCGCG 72
QY 72 CCCCCCTCTGAGGCGTCTTTCTTCAATTAAGAACATGGTGAACGATTCACACA 131
Db 73 CCCCCCTCTGAGGCGTCTTTCTTCAATTAAGAACATGGTGAACGATTCACACA 132
QY 132 TTAGCTGATCAGTGAAGATGTCACACTGCTGTCCTTCTCTTCCCTCTTGACT 191
Db 133 TTAGCTGATCAGTGAAGATGTCACACTGCTGTCCTTCTCTTCCCTCTTGACT 192
QY 192 TGTCTCTTGGACAAACAAATTCGCTGTACTGTCAGTGAAGTGAAGTGAAGTGAAGTGA 251
Db 193 TGTCTCTTGGACAAACAAATTCGCTGTACTGTCAGTGAAGTGAAGTGAAGTGAAGTGA 252
QY 252 TCTCAATGAAGTGTCTGTCATGTCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 311
Db 253 TCTCAATGAAGTGTCTGTCATGTCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 312
QY 312 ATTTGGCATGCTGTCACAGATGTCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 371
Db 313 ATTTGGCATGCTGTCACAGATGTCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 372
QY 372 ATGCTGCGAGTGAAGACAGCTGTCAGGCGCTGTGAGGCTTCCCACTTTCC 430
Db 373 ATGCTGCGAGTGAAGACAGCTGTCAGGCGCTGTGAGGCTTCCCACTTTCC 432
QY 431 AGACTCCACGCTGTTGGCATCAGGGGACGTCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 490
Db 433 AGACTCCACGCTGTTGGCATCAGGGGACGTCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 492
QY 491 GTCATCAAAATTAATAGATGTCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 550
Db 493 GTCATCAAAATTAATAGATGTCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 552
QY 551 TTCTCCAAATGAAGCTTCTTGTCTCACTGCTCTCTCATGTCGTCATTAAGTGAAGTGAAGTGAAGTGA 610
Db 553 TTCTCCAAATGAAGCTTCTTGTCTCACTGCTCTCTCATGTCGTCATTAAGTGAAGTGAAGTGAAGTGA 612
QY 611 TGATTAATGAAGTGTCTGTCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 670
Db 613 TGATTAATGAAGTGTCTGTCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 672
QY 671 TTTTCTTCAAGCAGCTTCTGATGTCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 730

Db	673	TTTTTCTTCAAGCAGCATGTTTCTGATGAGGAACAAGCTTTCAGTTTTCGACTGGCATC	732
Qy	731	ATGTGTCAGGATTTGCCAAGTCAAAATTTGGATTTGTTTCTTTTACCAATACCTTA	785
Db	733	ATGTGTTAGGATTTGCCAAGTCAAAATTTGGATTTGTTTCTTTTACCAATACCTTA	787
RESULT 5			
LOCUS	BMS43484	1010 bp	mRNA
DEFINITION	AGENCOURT 6492519 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5726572	linear	EST 20-FEB-2002
ACCESSION	BMS43484	5', mRNA sequence.	
VERSION	BMS43484.1	GI:18773915	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	1 (bases 1 to 1010)		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/ .		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgabbs@mail.nih.gov		
	Tissue Procurement: Invitrogen		
	cDNA Library Preparation: Life Technologies, Inc.		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)		
	DNA Sequencing by: Agencourt Bioscience Corporation		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNLN at:		
	http://image.lnl.gov		
	Plate: LHAM12719 row: e column: 05		
	High quality sequence start: 6		
	High quality sequence stop: 669.		
FEATURES			
source	Location/Qualifiers		
	1..1010		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5726572"		
	/tissue_type="hippocampus"		
	/lab_host="DH10B"		
	/clone_idb="NIH_MGC_124"		
	/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: EcoRV		
	(destroyed); Site 2: NotI; RNA source male hippocampus,		
	age 27. Library is oligo-dT primed and directionally		
	cloned (EcoRV site is destroyed upon cloning). Average		
	insert size 1.4 kb, insert size range 0.9-4 kb. Library is		
	normalized and enriched for full-length clones and was		
	constructed by C. Gruber (Invitrogen). Research Genetics		
	tracking code 012."		
ORIGIN			
Query Match	56.3%; Score 716.6; DB 4; Length 1010;		
Best Local Similarity	97.6%; Pred. No. 7.9e-194;		
Matches 760; Conservative	0; Mismatches 14; Indels 5; Gaps 3;		
Qy	12	TCACCTGGGGGCGGACAGTGAACCCGACCCCGCTGTGGGACCTTGAAGGGGAGATCCCGCGG	71
Db	87	TCACCTGGGCGGCGACGTGACCCGACCGCCGCTGGGACCTTGAAGGGGAGATCCCGCGG	146
Qy	72	CCCCCGCTCCCGACAGGCGTGTCTTTCTTCAATTAAGAACATGGTGAACATGATTCACACA	131
Db	147	CCCCCGCTCCCGACAGGCGTGTCTTTCTTCAATTAAGAACATGGTGAACATGATTCACACA	206
Qy	132	TTAGCTGATCAATGGAAGATGCAATGCTGCTGTGCTTCTCTCTTCCCTTGGCTACT	191
Db	207	TTAGCTGATCAATGGAAGATGCAATGCTGCTGTGCTTCTCTCTTCCCTTGGCTACT	266
Qy	192	TGCTTCCTTGGACAAACAAATTCGCTGTAATCTGTTAAGTGAATTAAGTGAACATGGCAAT	251
Db	267	TGCTTCCTTGGACAAACAAATTCGCTGTAATCTGTTAAGTGAATTAAGTGAACATGGCAAT	326

Qy	252	TTCTCATTGGAAGTTTCATCTACCTATGCTGTCCAGCTGCTGCTTTCTTCCCTTCAGGACAT	311
Db	327	TTCTCATTGGAAGTTTCATCTACCTATGCTGTCCAGCTGCTGCTTTCTTCCCTTCAGGACAT	386
Qy	312	ATTTTGGCAGTCGTGTGAACAGATGATGATACCACTGTCTCTATGGAAATACTGAAAAATGACAG	371
Db	387	ATTTTGGCAGTCGTGTGAACAGATGATGATACCACTGTCTCTATGGAAATACTGAAAAATGACAG	446
Qy	372	ATGCTGGCAGTGTATGAAACAGCCTTAGTGGCAGCCCTGTGAGGGTTTGGCAGTTTCCCA	431
Db	447	ATGCTGGCAGTGTATGAAACAGCCTTAGTGGCAGCCCTGTGAGGGTTTGGCAGTTTCCCA	506
Qy	432	GACCTCCAGCTGTTTGGCATCAGGGGGCAGCTGATGGAATGTGTTTGTGGAAATGACAG	491
Db	507	GACCTCCAGCTGTTTGGCATCAGGGGGCAGCTGATGGAATGTGTTTGTGGAAATGACAG	566
Qy	492	TCATACCAAAATATATATAGATGTGTATGATGTATTAAGATGGCTCTTGGCGGCAATGTGCATTT	551
Db	567	TCATACCAAAATATATATAGATGTGTATGATGTATTAAGATGGCTCTTGGCGGCAATGTGCATTT	626
Qy	552	TTCTCCTTAATGGAAGCTTCTTTTGTGATCAGTGGCTCTCTCATGTGGTATTTAACAGTGTGGAT	611
Db	627	TTCTCCTTAATGGAAGCTTCTTTTGTGATCAGTGGCTCTCTCATGTGGTATTTAACAGTGTGGAT	686
Qy	612	GATTAATAATGAGGATGCTGCAATGTAAGTGAAGAAAGCA--CATGATCTTGGAAATTAACCTGCTGGCA	670
Db	687	GATTAATAATGAGGATGCTGCAATGTAAGTGAAGAAAGCAATGATCTTGGAAATTAACCTGCTGGCA	746
Qy	671	TTTTTCTTTCACAGCCAGCTTTCTGTATGAGAACAA--GGTCTTTCAGTTTTCGACTGGCA	728
Db	747	TTTTTCTTTCACAGCCAGCTTTCTGTATGAGAACAAAGTCTTTCAGTTTTCGAAATGGCA	806
Qy	729	TCATGTGGTCAAGGATTTGCCAAGTCAAAA--TTTGGATTTGTTCTTTTACCATATCTTA	785
Db	807	TCATGTGGTCAAGGATTTGCCAAGTCAAAAATTTGGAATTTGTTCTTTTACCATATCTTA	865
RESULT 6			
Bi603184	603249828P1	NIH_MGC_96 Homo sapiens	CDNA clone IMAGE:5301559 5',
LOCUS	Bi603184	mRNA sequence.	
DEFINITION	Bi603184	EST.	
ACCESSION	Bi603184.1	GI:15496123	
VERSION			
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	NIH-MGC http://mgc.nci.nih.gov/.		
AUTHORS	National Institutes of Health, Mammalian Gene Collection (MGC)		
TITLE	Unpublished (1999)		
JOURNAL	Contact: Robert Strausberg, Ph.D.		
COMMENT	Email: cga@dbp-remail.nih.gov Tissue Procurement: Miklos Palkovits, M.D., Ph.D. CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki Toshiyuki and Piero Carninci (RIKEN) DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) CDNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.lnl.gov Place: LHAM11763 row: h column: 08 High quality sequence stop: 756. Location/Qualifiers		
FEATURES	1..858		
source	/organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:5301559" /tissue_type="hypothalamus"		

QY 312 ATTTGGCATGTGTTCAACAGATGTACCACTGTCTTATGAACTACTGAAATGACAG 371
| | | | |
DB 316 A-TTGGCATGTGTTCAACAGATGTACCACTGTCTTATGAACTACTGAAATGACAG 374
| | | | |
QY 372 ATGTGGCATGTGTTCAACAGATGTACCACTGTCTTATGAACTACTGAAATGACAG 431
| | | | |
DB 375 ATGTGGCATGTGTTCAACAGATGTACCACTGTCTTATGAACTACTGAAATGACAG 434
| | | | |
QY 432 GACTCCAGTGTGTTGGCATGAGGGGAGCTGATGAACTGTGTTTGTGGAATGACAG 491
| | | | |
DB 435 GACTCCAGTGTGTTGGCATGAGGGGAGCTGATGAACTGTGTTTGTGGAATGACAG 494
| | | | |
QY 492 TCATCAAAATTAATGATGTGTTGTTTAAAGATGCTCTTGGGAGCATGTGATTT 551
| | | | |
DB 495 TCATCAAAATTAATGATGTGTTGTTTAAAGATGCTCTTGGGAGCATGTGATTT 554
| | | | |
QY 552 TCTCTTAATGAGAGCTTCTTGTGATGCTGCTCTTC-ATGTGTTATTTTAAAGATGTTGGA 610
| | | | |
DB 555 TCTCTTAATGAGAGCTTCTTGTGATGCTGCTCTTC-ATGTGTTATTTTAAAGATGTTGGA 614
| | | | |
QY 611 TGATTAATGAGAGTGTGCTGATGTGAA-AAAGCATGATCTTGAATTAAC-TGCTGC 668
| | | | |
DB 615 TGATTAATGAGAGTGTGCTGATGTGAA-AAAGCATGATCTTGAATTAAC-TGCTGC 674
| | | | |
QY 669 GATTTTCTTCAACAGCAGTTTC-TGATGAGAAACAAGTCTTCAAGTTTTCGAC-TGG 726
| | | | |
DB 675 GATTTTCTTCAACAGCAGTTTC-TGATGAGAAACAAGTCTTCAAGTTTTCGAC-TGG 734
| | | | |
QY 727 CATCATGT-GGTCAAGATGTCAGTCA-NAITTTGATGTTTCTTTTACCATAT 781
| | | | |
DB 735 CATCATGTGGGTCAAGATGTCAGTCA-NAITTTGATGTTTCTTTTACCATAT 791
| | | | |

RESULT 8
CN793033 728 bp mRNA linear EST 26-MAY-2004
LOCUS 4127990 BARC 8BOV Bos taurus cDNA clone 8BOV_46H02 5', mRNA
DEFINITION
sequence.
ACCESSION CN793033
VERSION CN793033.1 GI:47689013
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
REFERENCE 1 (bases 1 to 728)
AUTHORS Baumann, R.G., Baldwin, R.L., Sonstegard, T.S., Van Tassel, C.P. and Matukumalli, L.K.
TITLE Construction and Analysis of a cDNA library Generated From Intestinal Muscle and Epithelial Tissues of Holstein Cattle
JOURNAL Unpublished (2004)
COMMENT Contact: Richard G. Baumann
Bovine Functional Genomics Lab
ARRI
BUDG 162: BARC-EAST, Beltsville, MD 20705, USA
Tel: 3015048604
Fax: 3015048744
Email: rbaumann@anri.barc.usda.gov
Single pass sequencing. Bases called and trimmed with phred 0.000925 using options -trim alt '- -trim fastavector identified by cross_match using options -mismatch 12 -mismatch 12
Plate: 46 row: H column: 02
Seq primer: CTATTTAGTGACATATAGAAC
High quality sequence stop: 728.
Location/Qualifiers
1..728
/organism="Bos taurus"
/mol_type="mRNA"
/strain="Holstein"
/db_xref="taxon:9913"
/clone="8BOV_46H02"
/sex="Female"

/tissue_type="Epithelial, Muscle"
/dev_stage="Lactating, Neonatal"
/lab_host="DH10B TONa"
/clone_lib="BARC 8BOV"
/notes="Organ: Intestine; Vector: pCMVSPORT6.1; Site_1: NCI; Site_2: EORI; Normalized cow cDNA intestinal library in pCMVSPORT6.1, constructed from equimolar mRNA pools derived from 5 sources, 4 lactating intestinal, 1 neonatal intestinal 4/5 lactating, proximal duodenum, jejunal, distal ileum, colon, 1/5 neonatal, proximal duodenum, jejunal, distal ileum"

ORIGIN

Query Match 41.5%; Score 527.8; DB 7; Length 728;
Best Local Similarity 88.2%; Pred. No. 1,1e-139;
Matches 574; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
QY 87 GCTGTTTCTTCAAAATTAAGAACATGTGAAATGATTCACATTAAGTATGATGT 146
| | | | |
DB 78 GCAGATATTTTCCAGTAAGAACATGTGAAATGATTCACATTAAGTATGATGT 137
| | | | |
QY 147 GAGATGTCACTGCTGCTGCTTCTCTTCCCTCTTGCTACTGCTCTTGACAA 206
| | | | |
DB 138 GATGATGTCAACGCTGCTGCTTCTCTTCCCTCTTGCTACTGCTCTTGACAA 197
| | | | |
QY 207 ACAATTCGCTTACTGCTGCTGCTTCTCTTCCCTCTTGCTACTGCTCTTGACAA 266
| | | | |
DB 198 ACAATTCGCTTACTGCTGCTGCTTCTCTTCCCTCTTGCTACTGCTCTTGACAA 257
| | | | |
QY 267 CATACCATGCTGCTGCTGCTGCTTCTCTTCCCTCTTGCTACTGCTCTTGACAA 326
| | | | |
DB 258 CACACCATGCTGCTGCTGCTGCTTCTCTTCCCTCTTGCTACTGCTCTTGACAA 317
| | | | |
QY 327 TCAACAGATGTACCACTGCTGCTGCTTCTCTTCCCTCTTGCTACTGCTCTTGACAA 386
| | | | |
DB 318 TCAACAGATGTACCACTGCTGCTGCTTCTCTTCCCTCTTGCTACTGCTCTTGACAA 377
| | | | |
QY 387 GAAAGGCTGCTGCTGCTGCTGCTTCTCTTCCCTCTTGCTACTGCTCTTGACAA 446
| | | | |
DB 378 GAAAGGCTGCTGCTGCTGCTGCTTCTCTTCCCTCTTGCTACTGCTCTTGACAA 437
| | | | |
QY 447 GCATCAGGGGAGCTGCTGCTGCTGCTTCTCTTCCCTCTTGCTACTGCTCTTGACAA 506
| | | | |
DB 438 GTCACAGGGGAGCTGCTGCTGCTGCTTCTCTTCCCTCTTGCTACTGCTCTTGACAA 497
| | | | |
QY 507 AGATGCTGCTGCTGCTGCTGCTTCTCTTCCCTCTTGCTACTGCTCTTGACAA 566
| | | | |
DB 498 AGATGCTGCTGCTGCTGCTGCTTCTCTTCCCTCTTGCTACTGCTCTTGACAA 557
| | | | |
QY 567 TTTCTTGTCACTGCTGCTGCTGCTGCTTCTCTTCCCTCTTGCTACTGCTCTTGACAA 626
| | | | |
DB 558 CTCCTTGTCACTGCTGCTGCTGCTGCTTCTCTTCCCTCTTGCTACTGCTCTTGACAA 617
| | | | |
QY 627 CTCATATGTAAGAAAGACATGATCTTGAATTAAGTCTGCTGCTTCTTGACAGCA 686
| | | | |
DB 618 CTCATATGTAAGAAAGACATGATCTTGAATTAAGTCTGCTGCTTCTTGACAGCA 677
| | | | |
QY 687 GTTCTGATGAGAACAGGCTTCAAGTTTTCAGCTGCTGCTGCTTCTTGACAGCA 737
| | | | |
DB 678 GTTCTGATGAGAACAGGCTTCAAGTTTTCAGCTGCTGCTGCTTCTTGACAGCA 728
| | | | |

RESULT 9
A1189142 505 bp mRNA linear EST 13-OCT-1998
LOCUS A1189142
DEFINITION
q040404.x1 Soares placenta 80weeks 2NHP809W Homo sapiens cDNA
clone IMAGE:1722702 3' similar to SW:72D3 DROMB P49846
TRANSCRIPTION INITIATION FACTOR TFIID 85 KD SUBUNIT 1, mRNA

ACCESSION A1189142
VERSION A1189142.1 GI:3740351
KEYWORDS EST.
SOURCE Homo sapiens (human)

QY 476 TTGTGGAAATGCAAGTCATCATCAATTATATAGATGTGGTGTAAAGATGGCTCTT 535

Db 181 TTGTGGAAATGCAAGTCATCATCAATTATATAGATGTGGTGTAAAGATGGCTCTT 240

QY 536 GGGCGCATGTGCATTTTCTCCTAAATGGAAGCTCTTGTCACTGGCTCCTATGTGTGA 595

Db 241 GGGCGCATGTGCATTTTCTCCTAAATGGAAGCTCTTGTCACTGGCTCCTATGTGTGA 300

QY 596 TTTAACAGTGTGGAGTGAATAAATGAGGTGTCTGCATGTGAATAAGACATGATCTTGG 655

Db 301 TTTAACAGTGTGGAGTGAATAAATGAGGTGTCTGCATGTGAATAAAGACATGATCTTGG 360

QY 656 AATTAACTGCTGCAGATTTTCTTCAACAGCCAGTTTCATGTGAGAAACAAGTCTTCAAGT 715

Db 361 AATTAACTGCTGCAGATTTTCTTCAACAGCCAGTTTCATGTGAGAAACAAGTCTTCAAGT 420

QY 716 TTTTCGACTGGCATCATGTGGTCAGGATTTGCAAGTCAAAATTTGGATTTGTTCTTTTAC 775

Db 421 TTTTCGACTGGCATCATGTGGTCAGGATTTGCAAGTCAAAATTTGGATTTGTTCTTTTAC 480

QY 776 CCATATCTTAG 786

Db 481 CCATATCTTAG 491

RESULT	11
LOCUS	B1759505
DEFINITION	B1759505 696 bp mRNA linear EST 25-SEP-2001 603046888BFL NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5187309 5' ,
ACCESSION	B1759505
VERSION	B1759505.1 GI:15751083
KEYWORDS	EST .
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 696) NIH-MGC http://mgc.nci.nih.gov/ . National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D.
AUTHORS	
TITLE	
JOURNAL	
COMMENT	

FEATURES
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: WGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <http://image.llnl.gov>
 Plate: LLM11467 row: 0 column: 22
 High quality sequence stop: 656.
 Location/Qualifiers

```

FEATURES
    source
        Location/Qualifiers
            1. .696
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="IMAGE:5187309"
               /lab_host="DH10B"
               /clone_11b="NIH MGC 116"
               /note="Organ: pooled colon, kidney, stomach; Vector:
               pcwv-sports; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
               source anonymous pool of 3 colons, age 26 yo male, 49 yo
               female, 71 yo male colon; 46 yo male kidney, and pool of 2
               stomachs, 62 yo male and 70 yo female. Library is
               oligo-dT primed and directionally cloned (EcoRV site is
               destroyed upon cloning). Average insert size 1.4 kb,
               insert size range 1-3 kb. Library is normalized and
               enriched for full-length clones and was constructed by C.
               Gruber (Invitrogen). Research Genetics tracking code
               023. Note: this is a NIH MGC library."

```

Query Match	38.2%;	Score 486.2;	DB 4;	Length 696;
Best Local Similarity	99.4%;	Pred. No. 9.1e-128;		
Matches 488; Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0;

Qy	296	CTCCCTTCACGACATATTTTGGATGCTGTTCACAGATGATGACCACTGCTCCATGGA	355
Db	1	CGCTCCGACAGGACATATTTTGGATGCTGTTCACAGATGATGACCACTGCTCCATGGA	60
Qy	356	TACTGAATAATGACAGATGCTGCGAGTATGGAACAGCTTAGTGGACCCCTGTGAGGGT	415
Db	61	TACTGAATAATGACAGATGCTGCGAGTATGGAACAGCTTAGTGGACCCCTGTGAGGGT	120
Qy	416	TTGCGAGTTTTCCCGACACTCCAGCTTTTGGCATCAGGGGCACTGATGGAACCTG	475
Db	121	TTGCGAGTTTTCCCGACACTCCAGCTTTTGGCATCAGGGGCACTGATGGAACCTG	180
Qy	476	TTTGTGGAATGACAGCTCATCAAAATATATAGATGTGTATGTTAAAGATGGCTCTT	535
Db	181	TTTGTGGAATGACAGCTCATCAAAATATATAGATGTGTATGTTAAAGATGGCTCTT	240
Qy	536	GGGGGAGTGGCATTTTCTCCATATGGAAGCTTTTGTCACTGGCTCCATGATGAGTGA	595
Db	241	GGGGGAGTGGCATTTTCTCCATATGGAAGCTTTTGTCACTGGCTCCATGATGAGTGA	300
Qy	596	TTTACAGTGGGATGATAAAATGAGGTGCTGCATAGTGAATAAAGACATGATCTTTGG	655
Db	301	TTTACAGTGGGATGATAAAATGAGGTGCTGCATAGTGAATAAAGACATGATCTTTGG	360
Qy	656	AATTACTGCTGCGATTTTCTTCACAGCCAGTTTGTATGATGAGAACAGGCTTTCAGTT	715
Db	361	AATTACTGCTGCGATTTTCTTCACAGCCAGTTTGTATGATGAGAACAGGCTTTCAGTT	420
Qy	716	TTTTTCAGTGGCATCATGTGTCAAGATGTGCCAAGTCAAAATTGGATTTGTTCTTTTAC	775
Db	421	TTTTTCAGTGGCATCATGTGTCAAGATGTGCCAAGTCAAAATTGGATTTGTTCTTTTAC	480
Qy	776	CCATATCTTAG 786	
Db	481	CCATATCTTAG 491	

RESULT 12	
AK011391	
LOCUS	2027 bp mRNA linear HTC 03-APR--2004
DEFINITION	Mus musculus 10 days embryo whole body cDNA, RIKEN full-length enriched library, clone:2610014F08 product:hypothetical SAM domain (sterile alpha motif)/Modified RING finger domain/G-protein beta WD-40 repeats containing protein, full insert sequence.
ACCESSION	AK011391
VERSION	AK011391.1 GI:12847483
KEYWORDS	HTC; CAP trapper.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	1
AUTHORS	Carninci, P. and Hayashizaki, Y.
TITLE	High-efficiency full-length cDNA cloning
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)
MEDLINE	99279253
PUBMED	10349636
REFERENCE	2
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itch, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE	20499374
PUBMED	11042159
REFERENCE	3
AUTHORS	Shibata, K., Itch, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., . . .

QY 787 ----- 786
Db 851 AATATAAAGACACTAAGTGGGACCTGCGCCCTGTTGGCTGTGCTTTTTCACATG 910
QY 787 ----- 786
Db 911 ATGGAAGATGCTTGATCGGGGTCAGTGATAATCTGTCATCATCATGATTCGACC 970
QY 787 ----- 786
Db 971 CTCAGAGTGTGCTACACACGCTGACTCAGCATACCAGTATGTTAGCATTTGCGTTG 1030
QY 787 ----- 786
Db 1031 CACCCAACTCTCTTACTTGTCTACTGTTCATATGACACAGACATGTAACATTTGGACG 1090
QY 787 ----- 825
Db 1091 TTGACCTGAAACACCTTGGCCAGAGCATGAGACCGCGTGAAACATTTGACTGAG 1150
QY 826 ATTGCTCAGAGAGTGTCTCAACATGCTTTGTGACACAGATTTAAAGATCTTTG 885
Db 1151 AATGCTCAGAGAGATGTCTCCGTGTGGCTTGTGCTCAAGGCTTGAAGACCTCGTGC 1210
QY 886 GTATTTCAAGATGATTAACATTTGAGAAAGAACTGTAATCTTACAAAGAAAGTC 945
Db 1211 GTATTTCAAGGCAACATTCATGAGAGAAAGAACTATGTCATCTCAACAGAAAGTC 1270
QY 946 TGCGCTGATGATTTGAAATGTAATCTCTAGAGCTGCGTAAAGTGTGAGAAATG 1005
Db 1271 TGCGTGTGATTTGAAATGTAATCTCTAGAGCTGCGAGAAAGTGTGAGAGATG 1330
QY 1006 AAGAGCTCAGAGCAAGGTTAATCTCTTTCTTGAAGATCTCTGATGATTTATATGTC 1065
Db 1331 AAGAGCTCAGAGCAAGGTTAATCTCTTTCTTGAAGATCTCTGATGATTTATATGTC 1390
QY 1066 CAATTAACAGAGCAATTAAGAAAGTCCGCTCATGCTGATGAGGCTATTCATATGAA 1125
Db 1391 CAATTAACAGAGCAATTAAGAAAGTCCGCTCATGCTGATGAGGCTATTCATATGAA 1450
QY 1126 AGAAGCAATGAAATTTGATTCAGCAAAAGAAAGTCAAGTCCCATGCAATCTTG 1185
Db 1451 GAGAAGCAATGAAATTTGATTCAGCAAAAGAAAGTCAAGTCCCATGCAATCTTG 1510
QY 1186 TTCTCTCTTACGCGTACTTACACCAATAGAGCTTGAATATGCGCATCATATGATGC 1245
Db 1511 CTCTCCCTTCACTGTACTGACCCCAAGACATGATGAGCATCAACGATGCGC 1570
QY 1246 TGGAGACACACCAAAATGAA 1266
Db 1571 TGGAGACACACCAAAATGAA 1591

RESULT 13
CN407189 463 bp mRNA linear EST 16-MAY-2004
LOCUS 1700600054632 GRN_PHEHP Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CN407189
VERSION CN407189.1 GI:47394734
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 463)
AUTHORS Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,
Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandilam, R.,
Lebkowski, J. and Stanton, L.W.
TITLE Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)
COMMENT Contact: Brandenberger R

Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 463 Std Error: 0.00.
Location/Qualifiers
1..463
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/issue_type="embryonic stem cells, DMSO-treated H9 cell
line"
/clone_1ib="GRN_PHEHP"
/note="Toigo dT primed, full-length enriched cDNA library
from DMSO-treated hES cell line H9 (p22) maintained in
feeder-free conditions"

ORIGIN
Query Match 36.4%; Score 463; DB 7; Length 463;
Best Local Similarity 100.0%; Pred. No. 3.6e-121; Indels 0; Gaps 0;
Matches 463; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 16 CTGGCGGCGACGTCGACCGCCGCTGGGACCTTGAAGCGGATCCCGCGGCC 75
Db 1 CTGGCGGCGACGTCGACCGCCGCTGGGACCTTGAAGCGGATCCCGCGGCC 60
QY 76 CGCTCTGACAGGCTGTTTTCTTCAATTAAGAAACATGTGAAATGATTCACATTA 135
Db 61 CGCTCTGACAGGCTGTTTTCTTCAATTAAGAAACATGTGAAATGATTCACATTA 120
QY 136 CTGATCATGTGACAGATGTCACATGCTGTCCTTCTCTTCCCTTGGCTACTGCT 195
Db 121 CTGATCATGTGACAGATGTCACATGCTGTCCTTCTCTTCCCTTGGCTACTGCT 180
QY 196 CTTTGACAAACAAATTCGCTGATCTGTTAGCTGATTCATGAACTGCAATCTC 255
Db 181 CTTTGACAAACAAATTCGCTGATCTGTTAGCTGATTCATGAACTGCAATCTC 240
QY 256 CATTGAAGTTTCACTAATCTGATCTGTCATGCTGCTGTTTTCTCCCTTCAGACATATTT 315
Db 241 CATTGAAGTTTCACTAATCTGATCTGTCATGCTGCTGTTTTCTCCCTTCAGACATATTT 300
QY 316 TGGCATGTGTTCAACAGATGTCACATGCTGCTTATGAAATGCAAAATGACATGTC 375
Db 301 TGGCATGTGTTCAACAGATGTCACATGCTGCTTATGAAATGCAAAATGACATGTC 360
QY 376 TGGCAGATGAGAACAGCTAGTGGCAGCCCTGTGAGGGTTTGCCAGTTTTCCCAAGCT 435
Db 361 TGGCAGATGAGAACAGCTAGTGGCAGCCCTGTGAGGGTTTGCCAGTTTTCCCAAGCT 420
QY 436 CCAAGTGTGTTGCAATCAGGGGCAAGCTGATGAACTGTGTTTT 478
Db 421 CCAAGTGTGTTGCAATCAGGGGCAAGCTGATGAACTGTGTTTT 463

RESULT 14
AL533462/c 967 bp mRNA linear EST 24-MAR-2004
LOCUS AL533462 Homo sapiens ADULT BRAIN Homo sapiens cDNA clone
DEFINITION CS00N004Y015 3-PRIME, mRNA sequence.
ACCESSION AL533462
VERSION AL533462.2 GI:31260543
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 967)
AUTHORS Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)


```

Db      420 TAGCTCATGTGCGCTGTGCGTTTCTCCGATGAGGCCCTTGTGCACTGGCTCCTC 479
Qy      587 ATGTGTGATTTTAAACAGTGTGGATGATPAAATGAGGTGTCTGATGATAAAGACA 646
Db      480 GGGCGGGAGCTTGACAGTGTGGATGACAGATGAGGTGTCTACACAGCGAAGGCGCA 539
Qy      647 TGATCTTGAATTAAGTGTGCGATTTTCTTCAACGCCAGTTTCTGATGAGAACAGG 706
Db      540 CGATCTGGGATGACCTGTGCTGACGCTTTCTCTCAGAGCTCTCTGCGGAG--AAGG 596
Qy      707 TCTTCAAGTTTTTTCAGCTGCGATGATGTGTGAGATTGCCAAGTCAAAATTGATTTG 766
Db      597 CTTCCAGCTTACACAGTTGGGCGTATGTGTCAAGACTGTGAATCAAACTGTGGCTGT 656
Qy      767 TTCTTTTACCCATATCTTA 785
Db      657 TACTATTACCGGTGTCTTA 675

```

Search completed: February 5, 2005, 08:18:30
 Job time : 4603 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 5, 2005, 10:22:21 ; Search time 5682 Seconds
(without alignments)
3274.693 Million cell updates/sec

Title: US-10-077-111-13

Perfect score: 2047
Sequence: 1 MVXLIHTLADHGDDVNCAP.....LTPNRTIKMAINRWLETHOK 384

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 4708233 segs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=x1h
-Q=/cgmt1/USFPO.spool/US10077111/runat_04022005_121052_6802/app_query.fasta.1.553
-DB=GenEmbl -QFMT=faaap -SUFPIX=p2n.rge -MINMATCH=0.1 -LOOPEC=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blousum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USR=US10077111.@CGN 1.1.4200.@runat_04022005_121052_6802 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEOBERRY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb_ba.*
2: gb_ncg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_ets.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2031	99.2	1564	9	BC029520 Homo sapi
2	1984	96.9	1996	6	CQ491312 Sequence
3	1984	96.9	1996	6	CQ497179 Sequence
4	1978	96.6	1811	6	AX748164 Sequence

5	1978	96.6	1811	9	AK093494
6	1978	96.6	1844	6	AK339277
7	1978	96.6	1919	9	AK129983
8	1823.5	89.1	1254	6	C0729621
9	1617	79.0	2130	10	BC061948
10	1464	71.5	3818	5	AJ719461
11	1373	67.1	1547	10	BC050792
12	1226.5	59.9	1291	6	AX076752
13	733	35.8	147021	2	AC006501
14	733	35.7	245795	9	AC096419
15	731	35.7	315463	2	AC122070
16	728	35.6	177089	9	AC008277
17	677	33.1	196403	2	AC102767
18	677	33.1	209437	2	AC111119
19	677	33.1	217352	10	AL929242
20	627.5	30.7	446	6	C0473162
21	626.5	30.6	409	6	C0482331
22	621.5	30.4	441	6	CQ503452
23	621.5	30.4	441	6	CQ512287
24	597	29.2	366	6	CQ096643
25	597	29.2	366	6	C0135438
26	597	29.2	366	6	CQ218732
27	597	29.2	366	6	CQ294537
28	597	29.2	366	6	CQ331208
29	552	27.0	2279	3	AK173854
30	532	26.0	297	6	C0109842
31	532	26.0	297	6	C0148540
32	532	26.0	297	6	CQ231810
33	532	26.0	297	6	CQ307124
34	532	26.0	297	6	CQ344013
35	318	15.5	399	6	AX071746
36	311	15.2	153866	9	AC009307
37	308.5	15.1	466	6	CQ097537
38	308.5	15.1	466	6	CQ136375
39	308.5	15.1	466	6	CQ219679
40	308.5	15.1	466	6	CQ258151
41	308.5	15.1	466	6	CQ295485
42	308.5	15.1	466	6	CQ332079
43	300	14.7	180	6	CQ110706
44	300	14.7	180	6	CQ149443
45	300	14.7	180	6	CQ232723

ALIGNMENTS

RESULT 1
BC029520
LOCUS
DEFINITION
Homo sapiens WD repeat and SAM domain containing 1, mRNA (cDNA clone MGC:33855 IMAGE:530155), complete cds.
ACCESSION
BC029520
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
REFERENCE
AUTHORS
Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Shat N.K., Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.T., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stadelton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulhavy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.W., Gay L.J., Hultk S.W., Villalón D.R., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D.,

Db	1226	CCTTAGCGGACTTACACCAAAATGAGACTCGAAAAATGGCCATCAATGATGCTGGAG	1285
QY	381	ThrHisGlnIyr 384	
Db	1286	ACACACCAAAAG 1297	
RESULT 2			
LOCUS	CQ491312	1996 bp	DNA
DEFINITION	Sequence 23179 from Patent WO0160860.		linear
ACCESSION	CQ491312		
VERSION	CQ491312.1	GI:41456931	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.		
JOURNAL	Schlegel, R., Endege, W.O. and Monahan, J.E.		
FEATURES	Genes differentially expressed in human prostate cancer and their use		
source	Patent: WO 0160860-A 23179, 23-AUG-2001;		
	Millennium Predictive Medicine, Inc. (US)		
	Location/Qualifiers		
	1..1996		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
Alignment Scores:			
Pred. No.:	1,966-172	Length:	1996
Score:	1984.00	Matches:	383
Percent Similarity:	80.46%	Conservative:	0
Best Local Similarity:	80.46%	Mismatches:	1
Query Match:	96.92%	Indels:	92
DB:	6	Gaps:	1
US-10-077-111-13 (1-384) x CQ491312 (1-1996)			
QY	1	MetValIyrVLeuIIeHisThiLeuAlaAspHisGlyAspAspValaAsnCysGylAAsp	20
Db	16	ATGGGAACCGATTACACACATTACTATCATGTGACGATGTCACATGCTGTGCTTC	75
QY	21	SerPheSerLeuLeuAlaThrCysSerLeuAspIyrStrIleArgLeuTySerLeuArg	40
Db	76	TCCTTTCCCTCTTGAGCTACTTGTCTCTTGGACAAACAAATTCGCTGACTCGTTACGT	135
QY	41	AspPheThrGluLeuProHisSerProLeuIyrPheHisThrTyAlaValHisCysCys	60
Db	136	GACCTTACTGAACTCCACATTCCTCCATTGAAGTTTCACTTAAGTCTGCACCTGCTGC	195
QY	61	CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrThrValIeu	80
Db	196	TGTTCTCTCCCTTCAGACATATTTTGGCATGTGTTCAACAGATGTACACACTGTCTTA	255
QY	81	TrpAsnThrGluAsnGlyGlnMetLeuAlaValMetGluGlnProSerGlySerProVal	100
Db	256	TGGAATACTGAAAATGAGACAGATGTGGCAGTGAAGAACAGCCTAATGGGACCCCTGTG	315
QY	101	ArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyValAlaAspGlyThr	120
Db	316	AGGGTTTCCCACTTTTCCCCAGACTCCACGTTGTTGGCATCAAGGGGACGCTGATGAAC	375
QY	121	ValValLeuTrpAsnAlaGlnSerTyIyrLeuTyArgCysGlySerValIyrAspGly	140
Db	376	GTGGTTTGTGGAATGACACAGTCATACAAATATATATAGTGTGATGTTAAAGATGGC	435
QY	141	SerLeuAlaAlaCysAlaPheSerProAsnGlySerPhePheValThrGlySerSerCys	160
Db	436	TCTTTGGGGGACATGCAATTTTCTCTTAATGGAACCTCTTGTCACTGGCTCCCTCAATG	495

OY		161	GlyAspLeuThrValLTPAspAspLysMetArgCysLeuHISerGluValAlaHisasp	180
Db		496	GGTGTATTAACAGCTGGCATGATTAATAATGAAGTGCTGCATAGTAAAAAAGCCATGAT	555
OY		181	LeuGIyIleTrnCysCyAspPheSerSerInProValSerAspGlygluGlnGlyeu	200
Db		556	CTTGGAATTACTCTGCTCGGATTTTTCCTTCACAGCCAGTTCTGATGAGAACAAGCTCTT	615
OY		201	GlnPhePheArgLeuAlaSerCysGlygluAspCysGlnValLysIleTrpIleValSer	220
Db		616	CAGTTTTTTCAGCTGGCATCATGTGTGTCAGGATTTGCCAAGTCAAAATTGGATTGTTCT	675
OY		221	PheThrHISileu-----	225
Db		676	TTTACCACATATCTTAGGTTTGAATTAATAATATAAGTACATGAGTGGCACTGTGCT	735
OY		225	-----	225
Db		736	CCTGTTCTGCTGTGTGCTTTTCCCAGATGGCAGATGCTAIGTCTCAGGCTCAGTGAT	795
OY		225	-----	225
Db		796	AAGTCTGTCAATAGTATAATGATACTAATACTAGAAATAATCTTCACACATGACTCAGCAC	855
OY		225	-----	225
Db		856	ACCAAGTATGTCAAACTGTGTCTTTGGACCTTAATACCTTTTACTTGGCTACTGTGTCA	915
OY		226	-----AlaArgArg	228
Db		916	ATGACAAACACAGTGAACATCTGGCAATTGACCTGGAACACTTTGGCCAAGCAAGGGC	975
OY		229	ThrgluHISglnLeuLysGlnPheThrGluAspTrpSerGluGluValIleSerThrTrp	248
Db		976	ACAGAACCTCAGCTGAGCAATTTTACCGAAGATTGTCTAGAGAGAGATGTCTCAACATGG	1035S
OY		249	LeuCysAlaglnAspLeuLysAspLeuValGlyIlePheLysMetAsnAsnIleAspGly	268
Db		1036	CTTTGTGCACAAGATTTTAAAGATCTTGTTGGTATTTTCAAGATGAAATPAACATTGATGGA	1095S
OY		269	LysGluLeuLeuAsnLeuThrLysGluSerLeuAlaAspAspLeuValIleGluSerLeu	288
Db		1096	AAAGAACCTGTGAATCTTACAAAGAAAGTCGTGCTGATGATTTGAAATTTGAATCTCTTA	1155S
OY		289	GlyLeuArgSerLysValIleuArgLysIleGluGluLeuArgThrLysValLysSerLeu	308
Db		1156	GGACTCGCTAGTAAGTCTGAGGAANAATTAAAGAGCTCAGACACCAAGTTAAATCCCTT	1215S
OY		309	SerSerGlyIleProAspGluPheIleCysBroIleThrArgGluLeuMetLysAspPro	328
Db		1216	TCTTCAGGAATTCCTGATGATTTTATATGTCCATTACTAGAAACCTTATGAAGAATCCG	1275S
OY		329	ValIleAlaSerAspGlyLysTrpSerTyrluLysGluAlaMetGluAsnTrpIleSerLys	348
Db		1276	GTCATCGCATCAGATGGCTATTTCATATGAAAGAAAGACAAATGGATAACGCAAA	1335S
OY		349	LysLysArgTrpAspPrometThrAsnLeuValLeuProSerAlaValIleLeuThrProasn	368
Db		1336	AAGAAACGTACAAAGTCCCATCAAAATCTTGTTCTTCCTTAGGGGTATCTTAACCAAT	1395S
OY		369	ArgThrLeuLysMetAlaIleAsnArgTrpLeuGluLuthrHisGlnLys	384
Db		1396	AGGACTCTGAATAATGGCCATCAATAGATGCTGGAGACACACCAAAG	1443
RESULT 3				
LOCUS	CQ497179	1996 bp	DNA	linear PAT 30-JAN-2004
DEFINITION	Sequence 29046 from Patent WO0160860.			
ACCESSION	CQ497179			
VERSION	CQ497179.1	GI:41462815		
KEYWORDS	.			
ORGANISM	Homo sapiens (human)			
SOURCE	Homo sapiens			

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
1 Schlegel,R., Endege,W.O. and Monahan,J.E.
TITLE Genes differentially expressed in human prostate cancer and their use

JOURNAL
Patent: WO 0160860-A 29046 23-AUG-2001;
Millennium Predictive Medicine, Inc. (US)

FEATURES
Source 1. .1996
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 1.96e-172 Length: 1996
Score: 1984.00 Matches: 383
Percent Similarity: 80.46% Conservative: 0
Best Local Similarity: 80.46% Mismatches: 1
Query Match: 96.92% Indels: 92
DB: Gaps: 6

US-10-077-111-13 (1-384) x CQ497179 (1-1996)

QY 1 MetValIyLeuIleHsiHrLeuAlaAspHisGlyAspAspValaAsnCySyaAlaPhe 20
Db 16 ATGCTGAACCTGATTACACATTAGTCATGATGATGACGATGCACTGCTGCTTC 75
QY 21 SerPheSerLeuLeuIleThrCySerSerLeuAspIleThrIleArgLeuIleArg 40
Db 76 TCCCTTCCCTCTTGGCTACTGCTCTCTTGGACAAACAAATTCCTGCTGCTTACGT 135
QY 41 AspPheThrGluLeuProHisSerProLeuLeuPheHisThrTyraIleValHisCySya 60
Db 136 GACTTACTGAACTGCACACTTCTCCATTGAAGTTTCACTACTGCTGCTGCTGCTGC 195
QY 61 CyasPheSerProSerGlyHisIleLeuAlaSerCySerThrAspGlyThrThrValleu 80
Db 196 TGTTTCCCTCCCTCAGAGACATATTTGGCATCGTGTCAACAGATGGTACCACTGCTCTA 255
QY 81 TrpAsnThrGluLeuArgIleMetLeuAlaValMetGluIleProSerGlySerProVal 100
Db 256 TGGAAATACATAAATGACAGATGCTGGAGTGTGAGAACAGCCTTACTGACAGCCCTGTG 315
QY 101 ArgValCySerGlnPheSerProAspSerThrCySerLeuAlaSerGlyAlaIleAspGlyThr 120
Db 316 AGGCTTTGCCAGTTTCCCAAGACTCCAGCTGTTGGCATCAGGGGCACTGATGAACT 375
QY 121 ValValIleuTrpAsnAlaGlnSerTyrlsLeuIleuTyArgCySgIySerValIlyAspGly 140
Db 376 GTGGTTTGTGAATGACAGATGATACAAATTAATATGATGTGTGTTAAAGATGGC 435
QY 141 SerLeuAlaIleCySyaIlePheSerProAspGlySerPhePheValThrGlySerSerCyS 160
Db 436 TCCCTGGCGCATGTGCATTTTCTCCPAAATGGAAGCTTCTTGTGCTGCTGCTCTCATGT 495
QY 161 GlyAspLeuThrValIleTrpAspAspIleMetArgCySerLeuHisSerGlyIleValHisAsp 180
Db 496 GGTGATTTAACAGTGTGGATGATTAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 555
QY 181 LeuGlyIleThrCySyaAspPheSerSerGlnProValSerAspGlyGluGlnGlyLeu 200
Db 556 CTGGGAATTAACCTGCTGCTGCTGCTTCTTCAACGCACTTCTGATGAGAAACAAGCTT 615
QY 201 GlnPhePheArgLeuAlaSerCySgIyGlnAspCySgIleValIleTrpIleValSer 220
Db 616 CAGTTTTCGACTGGCATGTGTGTGAGGATGCAAGATGCAAAATTTGGATTGTTCT 675
QY 221 PheThrHisIleLeu----- 225
Db 676 TTTAACCAATATCTTAAGTTTGAATTAAATATATAAAGTACATGAGTGGGCACTGTGCT 735

QY 225 ----- 225
Db 736 CCTGTTCTGGCTTGTGCTTTTCCATGATGGCAGATGCTAGTCAAGGTGATGAT 795
QY 225 ----- 225
Db 796 AAGTCTGTCACTAGTATATGATTAATACTGAGAAATATACTTACACATTGACTCAGCAC 855
QY 225 ----- 225
Db 856 ACCAGGATGTCAACAATTGTGCTTTTGACACCTAATACCTTTTACTGCTACTGTTCA 915
QY 226 -----AlaArg 228
Db 916 ATGACAAACACAGTGAACATCTGCAATTGACCTGGAACACTTTGCCAACAGAGCCG 975
QY 229 ThrGluHisGlnLeuLeuGlnPheThrGluAspTrpSerGlnGluValIleSerThrTrp 248
Db 976 ACAGAACATCAGCTGGAAGCAATTTACCGAAGATTGTCAGAGAGGATGTCTCAACATGG 1035
QY 249 LeuCyAlaGlnAspLeuIleuValGlyIlePheIleMetAsnAsnIleAspGly 268
Db 1036 CTTTGTGCACAAATTTAAAGATCTTGTGGTATTTTCAAGTGAATTAACATGATGAGA 1095
QY 269 LysGluLeuLeuAsnLeuThrIlyGlnSerLeuAlaAspAspLeuIleGlnSerLeu 288
Db 1096 AAAGAACTGTTGAATCTTCAAAAGAAAGCTGCTGATGATTTGAAATTCCTTA 1155
QY 289 GlyLeuAspSerIleValLeuArgIleIleGluIleuAspThrIlyValIlySerLeu 308
Db 1156 GGACTGGTATGAAGTGTGAGGAATAATGAAAGCTCAGGACCAAGGTTAAATCCCTT 1215
QY 309 SerSerGlyIleProAspGluPheIleCyProIleThrArgGluLeuMetIlyAspPro 328
Db 1216 TCTTCAGGAATTCCTGATGATTAATATATGTCGAATACAGAACTTATGAAGATCCG 1275
QY 329 ValIleAlaSerAspGlyTyrlsSerTyrgIlyGlnIleMetGluAsnTrpIleSerIly 348
Db 1276 GTCATGCAATCAGATGCTATTCATATGAAAGAAAGCAATGAAATTTGATCAGCAAAA 1335
QY 349 LysIleArgThrSerProMetThrAsnLeuValleuProSerIleValleuTrpProAsn 368
Db 1336 AAGAAACGTACAACTCCACAGCAAAATCTTGTCTTCAAGCGGTACTTACACCAAAAT 1395
QY 369 ArgThrLeuIleMetAlaIleAsnArgTrpLeuGluThrHisGlnIly 384
Db 1396 AGGACTCTGAATAATGGCCATCAATATGATGCTGGAACACACCAAAAG 1443

RESULT 4
AX748164 1811 bp mRNA linear PAT 20-JUN-2003
LOCUS
DEFINITION
Sequence 1689 from Patent EP1308459.
ACCESSION
AX748164
VERSION
AX748164.1 GI:32132552
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
1 Isogai,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S., Yamamoto,J.I., Isono,Y., Hio,Y., Otsuka,K., Nagai,K., Irie,R., Tamechika,I., Seki,N., Yoshikawa,T., Otsuka,M., Nagahara,I.K. and Masuo,Y.
TITLE Full-length cDNA sequences
JOURNAL
Patent: EP 1308459-A 1689 07-MAY-2003;
Helix Research Institute (JP) ; Research Association for Biotechnology (JP)
FEATURES
Source 1. .1811
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 6,13e-172 Length: 1811
 Score: 1978.00 Matches: 382
 Percent Similarity: 80.25% Conservative: 0
 Best Local Similarity: 80.25% Mismatches: 2
 Query Match: 96.63% Indels: 92
 DB: 6 Gaps: 1

US-10-077-111-13 (1-384) x AX748164 (1-1811)

QY 1 MetValylsLeuIleHisThrLeuAlaAspHisGlyAspAspValaIasnCysCysAlaPhe 20
 Db 150 ATGCTGAACATGATTCACACATTGATGATCATGTCAGATGTCACATGCTGTGCGCTTC 209
 QY 21 SerPheSerLeuLeuAlaThrCysSerLeuAspLysThrIleArgLeuYrSerLeuArg 40
 Db 210 TCCCTTTCCCTCTGGCTACTTGTCTCTGGACAAACAAATCCGCTGACTCCTTACGT 269
 QY 41 AspPheThrGluLeuProHisSerProLeuLysPheHisThrYrAlaValHisCysCys 60
 Db 270 GACTTTACTGACGACCAATCTCCACATGAAGTTTCAATGCTGATGCTGCTGCTGCTG 329
 QY 61 CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrThrValLeu 80
 Db 330 TGTTCCTCCCTTCAGGACATATTTGGCATCGTGTTCACAGATGCTACCATCTGCTCTA 389
 QY 81 TrpAsnThrGluAsnGlyGlnMetLeuAlaValMetGluGlnProSerGlySerProVal 100
 Db 390 TGGAAATCTGAAAAATGACAGATCTGCGAGTGTAGTGAACAGCTTAGTGGCAGCCCTGTG 449
 QY 101 ArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyAlaAlaAspGlyThr 120
 Db 450 AGGCTTTGCCAGTTTCCCACTCCAGCTCCAGTGTGGCATCAAGGCGACGCTGATGAGCT 509
 QY 121 ValValLeuThrPheAsnAlaGlnSerTyrLysLeuYrArgCysGlySerValLysAspGly 140
 Db 510 GTGCTTTTGTGAGATGACAGCTCATCAAAATTAATAGATGTGATGTTAAAGATGGC 569
 QY 141 SerLeuAlaAlaCysAlaPheSerProAsnGlySerPheValIleArgLysSerCys 160
 Db 570 TCCCTGGCGGATGTGATTTTCTCTAATGAAAGCTTCTTGTGACATGGCTCTCTCATGT 629
 QY 161 GlyAspLeuThrValIleTrpAspAspLysMetArgCysLeuHisSerGlyLysValHisAsp 180
 Db 630 GGTGATTTAACAGTGTGGATGATMAATGAGGTGTCTGCAATGTGAAAAAGCACATGAT 689
 QY 181 LeuGlyIleThrCysCysAspPheSerSerGlnProValSerAspGlyGluGlnGlyLeu 200
 Db 690 CTTGGAAATTACCTCTGCGATTTTCTTCAAGCCAGTTTCTGATGGAAACAAAGGCTTT 749
 QY 201 GlnPhePheArgLeuAlaSerCysGlyGlnAspCysGlnValLysIleTrpIleValSer 220
 Db 750 CAGTTTGTGACATCGGATCATGTGTCAGAGATTGCCAAGCAAAATTTGGATTGTTCT 809
 QY 221 PheThrHisIleLeu----- 225
 Db 810 TTTTACCAATATCTTAGGTTTGAATTAATAATATAAGTACATGAGTGGGCACTGTGCT 869
 QY 225 ----- 225
 Db 870 CCTGTTCTGGCTGTGCTTTTCCCATGANGGCGAGATGCTAGTCTCAGGGTCAGTGAT 929
 QY 225 ----- 225
 Db 930 AAGTCTGTCATAGTATATGATTAATATCTGAGAATATATCTTCAACACATTGACTCAGAC 989
 QY 225 ----- 225
 Db 990 ACCAGTATGTCAACAATTGTGCTTTGGCACTTAATACCTTTTACTTGTACTGCTGTTCA 1049
 QY 226 -----AlaArgArg 228

Db 1050 ATGACAAACAAACAGTGAACATCTGGCAATTGACCTGGAAACACTTGGCCAAAGAGAC 1109
 QY 229 ThrGlnHisGlnLeuLysGlnPheThrGluAspTrpSerGlyGluValIleSerThrTrp 248
 Db 1110 ACGAATATAGAGCTGAAGAGCAATTTTCCAGAGATTTGTCAAGAGAGATGTCACAACTGG 1169
 QY 249 LeuCysAlaGlnAspLeuLysAspLeuValGlyIlePheLysMetAsnValIleAspGly 268
 Db 1170 CTTTGTGCACAAGATTTTAAAGATCTTGTGTGATTTTCAAGATGATTAATCATGATGGA 1229
 QY 269 LysGluLeuLeuLysLeuThrLysGluSerLeuAlaAspAspLeuLysIleGluSerLeu 288
 Db 1230 AAAGAACTGTTGAATCTTAAACAAAGAAAGCTGCTGATGATTTGAAAATTTGAATCTTA 1289
 QY 289 GlyLeuAspSerLysValIleuArgLysIleGluGlnLeuArgThrLysValLysSerLeu 308
 Db 1290 GGACTGCGTGAAGTGAAGCTGAGAGAAATTTGAAGAGCTCAGACCAAGTTTAAATCCCTT 1349
 QY 309 SerSerGlyIleProAspGluPheIleCysProIleThrArgGluLeuMetLysAspPro 328
 Db 1350 TCTTCAGGAATTTCTGATGATTAATTAATATGTCATTAATGAGAACTTAAGAAAGATCCG 1409
 QY 329 ValIleAlaSerAspGlyTyrSerTyrGluLysGluAlaMetGluAsnTrpIleSerLys 348
 Db 1410 GTATGTCATCAGATGCTGCTATTCATTAATGAAGAAAGAACATGAGAAATTTGATCAGCAA 1469
 QY 349 LysLysArgThrSerProMetThrAsnLeuValLeuProSerAlaValIleuThrProAsn 368
 Db 1470 AAGAAAGTGAACAGTCCCATGACAAACAAATCTTCTTCTTCCACGGTACTTAAACCAAT 1529
 QY 369 ArgThrLeuLysMetAlaIleAsnArgTrpLeuGlnThrHisGlnLys 384
 Db 1530 AGGACTGTGAATATGGCCATCAATAGTGGCTGGAGACACCAAAAG 1577

RESULT 5

AK093494

LOCUS

DEFINITION

Homo sapiens cDNA FLJ36175 fis, clone TEST12026453, weakly similar

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,
 Makamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H.,
 Sekine, M., Ohtsuka, M., Nishi, T., Shibahara, T., Tanaka, T.,
 Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,
 Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M.,
 Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,
 Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T.,
 Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsura, N., Sato, K.,
 Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yanashita, H.,
 Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M.,
 Hiraoaka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S.,
 Yoshida, M., Hottu, T., Kusano, J., Kanehori, K., Takahashi, Fujii, A.,
 Hara, H., Tanase, T., Nomura, Y., Togaya, S., Komai, F., Hara, R.,
 Takeuchi, K., Arita, M., Imose, N., Kusashino, K., Yuki, H., Oshima, A.,
 Sasaki, N., Aotsuma, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T.,
 Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S.,
 Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,
 Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,
 Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,
 Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M.,
 Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yanada, K., Fujii, Y.,
 Ozaki, K., Hirose, M., Omori, Y., Kawabata, A., Hikiji, T., Kobatake, N.,
 Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T.,
 Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsunura, K.,
 Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T.,

TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J.,
Sato, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K.,
Nagase, T., Nomura, N., Kikuchi, H., Masuo, Y., Yamashita, R.,
Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isegai, T. and Sugano, S.
Complete sequencing and characterization of 21,243 full-length
human cDNAs
Nat. Genet. 36 (1), 40-45 (2004)
14702039

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (04-JUL-2002) Takao Isegai, FLJ Project (HRI Team); 2-6-7
Kazuma-Kamata, Kibatazu, Chiba 292-0612, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
Location/Qualifiers
1. 1811
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BEST12026453"
/issue_type="test18"
/clone_lib="BEST12"
/note="cloning vector: pME18SFL3"
150. .1580
/note="unnamed protein product"
/cdon_start=1
/protein_id="BAC04184.1"
/db_xref="GI:21752385"
/translation="MVKLIHTLADHDGDVNCASFPSGLATGSLDKTIRLYSLDFTE
LPHSPKFTYAVHCCFSPSGHILASCTDGTIVLMTNGQMLVMVEQPSGPRV
CQSPSPSTCLASGADGTVYLMNAQSYKLYRCGSVDGSLAACAFSPNGSPVYTGSS
GDLVMDDKMRCKHSEKADHDLITCCPFSQPSVDEGQLOFRLASCGDDCVKXIM
VSFTHLIGPELKYKRTLSGCHAPVLCASFSDQMLVBSGVDSVIVYDPTNLIHT
LTOHTRYVTTCAFPNTLLATSGMDKTIVIKQFDSTLCOASTHQLQFTEDVSE
EDVSTWCAQDLKDLVGIFFPMNIDGELNLTKESTLADLKLKESGLRSKVLAKIE
LRTKXVSLSSGIDPEFCIPTRLEMKDPVLASDGYSEKEMENWISKKRTSPMTNL
VLPSAVLTPRRLKMAINRWLETHOK"

COMMENT
FEATURES
source
1. 1811
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BEST12026453"
/issue_type="test18"
/clone_lib="BEST12"
/note="cloning vector: pME18SFL3"
150. .1580
/note="unnamed protein product"
/cdon_start=1
/protein_id="BAC04184.1"
/db_xref="GI:21752385"
/translation="MVKLIHTLADHDGDVNCASFPSGLATGSLDKTIRLYSLDFTE
LPHSPKFTYAVHCCFSPSGHILASCTDGTIVLMTNGQMLVMVEQPSGPRV
CQSPSPSTCLASGADGTVYLMNAQSYKLYRCGSVDGSLAACAFSPNGSPVYTGSS
GDLVMDDKMRCKHSEKADHDLITCCPFSQPSVDEGQLOFRLASCGDDCVKXIM
VSFTHLIGPELKYKRTLSGCHAPVLCASFSDQMLVBSGVDSVIVYDPTNLIHT
LTOHTRYVTTCAFPNTLLATSGMDKTIVIKQFDSTLCOASTHQLQFTEDVSE
EDVSTWCAQDLKDLVGIFFPMNIDGELNLTKESTLADLKLKESGLRSKVLAKIE
LRTKXVSLSSGIDPEFCIPTRLEMKDPVLASDGYSEKEMENWISKKRTSPMTNL
VLPSAVLTPRRLKMAINRWLETHOK"

ORIGIN
Alignment Scores:
Pred. No.: 6,13e-172 Length: 1811
Score: 1978.00 Matches: 382
Percent Similarity: 80.25% Conservative: 0
Best Local Similarity: 80.25% Mismatches: 2
Query Match: 96.63% Indels: 92
DB: 9 Gaps: 1
US-10-077-111-13 (1-384) x AK093494 (1-1811)

QY 1 MewVallyleuileuilehsthrleuAlaAspHieGlyAspAspValaenCyseValaIaphe 20
DB 150 ATGGTGAAGCTGATTCACACATTAGCTGATCATGGTGAAGATGCAATGCTGCTGCTTC 209

QY 21 SerPheSerleuLeuAlaThrCyseSerleuAspIlyThrIleArgIeuIyrSerleuArg 40
DB 210 TCTTTTTCCTCTTGGCTACTGGCTCTTGGACAAAACAAATTCGCTGACTGTTACGT 269
QY 41 AspPheThrGluIeuProHisSerProleuIysPheHieThrTyraIaValHisCyseCy 60
DB 270 GACTTACTGAACTGGCACTTCTCCATTGAACTTCACTGATCTGCTCCCTGCTG 329
QY 61 CyseSerProSerGlyHisIleleuAlaSerCyseSerThrAspGlyThrThraValleu 80
DB 330 TGTTCCTCCCTTCAGACATATTTGGCATGGTGTTCACAGATGTAACACTGTCCTA 389
QY 81 TrpAsnThrGluAsnGlyIleMetleuAlaValMetGlnIleProSerGlySerProVal 100
DB 390 TGGAAATACGAAATAGACAGATGCTGGCAGATGGAACAGCTGTGTGGCCGCTGTG 449
QY 101 ArgValCyseGlnPheSerProAspSerThrCyseleuAlaSerGlyIaIaAspGlyThr 120
DB 450 AGGGTTTGGCAGTTTCCCGACATCCAGCTGTTGGCATCAGGGGACGCTGATGGAACT 509
QY 121 ValValleuIleTrpAsnIaGlnSerTyrllyleuTyArgCyseGlySerValIysAspGly 140
DB 510 GTGGTTTGTGGAAATGACAGTCAATCAATTAATATGATGTGTATTAAGATGGC 569
QY 141 SerleuAlaIaCyseAlaPheSerProAsnGlySerPhePheValThrGlySerSerCy 160
DB 570 TCTTGGCGGCGATGTCATTTTCTCTTAATGGAAGCTTCTTGTCACTGGCTCCTCATGT 629
QY 161 GlyAspIleuThrValIleTrpAspIlyIleMetArgCyseleuHisSerGlyIysIaHisAsp 180
DB 630 GGTGATTTAAAGTGTGGATGTAATAAGAGTGTCTCATATGTAATAAGAACATGAT 689
QY 181 IeuGlyIleThrCyseCyseAspPheSerSerGlnProValSerAspGlyGluGlnGlyleu 200
DB 690 CTGGAAATTAACGTCGTCGCAATTTTCTTCAAGCCAGCTTTCGATGGAACAGCTCTT 749
QY 201 GlnPhePheArgleuAlaSerCyseGlyIleAspCyseGlnValIysIleTrpIleValSer 220
DB 750 CAGTTTTCGACTGGCATCATGTGTGTCAGAGATTGCCAAGTCAAAATTTGGATTGTTCT 809
QY 221 PheThrHisIleIleu----- 225
DB 810 TTTAACCAATATCTTAGCTTTTGAATTAATAATAAATGAACACTGAGTGGCATGTGCT 869
QY 225 ----- 225
DB 870 CCTGTTCTGGCTTGTGCTTTTCCCATGATGGGACAGATGCTACAGGTCATGAT 929
QY 225 ----- 225
DB 930 AAGTCTGATATGATATGATATTAATACTGAGAATAATTACTCACACATTGACTCAGAC 989
QY 225 ----- 225
DB 990 ACCAGGATATGCAACATTGTGCTTTTGACCTTAATACCTTTTACTTCTACTGCTTCA 1049
QY 226 ----- 228
DB 1050 ATGGACAAAACACTGAACATCTGGCAATTGACCTGGAAAACCTTTGCCAAGCAGAGAC 1109
QY 229 ThrGlnHisGlnIleuIysGlnPheThrGluAspTrpSerGlyIleValIaSerThrTrp 248
DB 1110 AACGAACATCAGCTGAAGCAATTTTACGAAGATTTGCTCAGAGAGAGATCTTCAACATGG 1169
QY 249 IeuCyseAlaGlnAspIleuIysAspIleuValGlyIlePheIysMetAsnAsnIleAspGly 268
DB 1170 CTTTGTGCAACAAGATTTAAAGATCTTGTGATTTTTCMAAGATGAATTAACATTGATGA 1229
QY 269 IysGluIleuIysAsnIleuThrIysGlnIleuAlaAspAspIleuIysIleGlnIleu 288
DB 1230 AAAGAACTGTGAATTTTCAAAAGAAAGTCTGGCTGAGATTTGAAAATTTGAATCTCTTA 1289
QY 289 GlyIeuArgSerIysValIleuArgIysIleGluGluIleuArgThrIysValIysSerIeu 308

AKI29983	1919 bp	mRNA	linear	PRI 10-SEP-2003
LOCUS	AKI29983			
DEFINITION	Homo sapiens CDNA FLJ26473 f1s, clone KDN04521.			
ACCESSION	AKI29983			
VERSION	AKI29983.1 GI:34526702			
KEYWORDS	oligo capping; f1s (full insert sequence).			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Euteleostomi; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
AUTHORS	1 Taehiro,H., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M., Suzuki,Y., Hata,H., Nakagawa,K., Mizuno,S., Morinaga,M., Kawamura,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Nishikawa,T., Sugiyama,A., Kawakami,B., Nagai,K., Isogai,T. and Sugano,S.			
TITLE	NEDO human cDNA sequencing project			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 1919)			
AUTHORS	Sugano,S. and Suzuki,Y.			
TITLE	Direct Submission			
JOURNAL	Submitted (31-JUN-2003) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail: flicdn@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)			
COMMENT	NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction and 5'-end and one pass sequencing: Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; 3'-end one pass sequencing: RAB; clone selection for full insert sequencing: RAB and Helix Research Institute.			
FEATURES	Location/Qualifiers			
source	1. 1919			
	/organism="Homo sapiens"			
	/mol_type="mRNA"			
	/db_xref="taxon:9606"			
	/clone="KDN04521"			
	/tissue_type="Kidney"			
	/clone_lib="KDN"			
	/note="cloning vector: PME18SFL3"			
ORIGIN				
Alignment Scores:				
Pred. No.:	6,62e-172	Length:	1919	
Score:	1978.00	Matches:	382	
Percent Similarity:	80.25%	Conservative:	0	
Best Local Similarity:	80.25%	Mismatches:	2	
Query Match:	96.63%	Indels:	92	
DB:	9	Gaps:	1	
US-10-077-111-13 (1-384) x AKI29983 (1-1919)				
QY	1 MetValLysLeuIleHisThrLeuAlaAspHisGlyAspAspValAsnCysCysAlaPhe	20		
DB	258 ATGGGAAACGATTCACACATTACTGATCATGCTGACGATGCAATGCTGCTGCTTC	317		
QY	21 SerPheSerLeuLeuAlaThrCysSerLeuAspLysTrpIleArgLeuTyrSerLeuLys	40		
DB	318 TCCCTTTCCCTTGCGCTACTGCTGCTTGGCAAAACAAATTCGCTGACTCGTTACGT	377		
QY	41 AspPheThrGluLeuProHisSerProLeuLysPheHisThrTyrAlaValHisCysCys	60		
DB	378 GACTTTACTGAACTGCGACATTCCTCATTTGAAGTTTCATTACTTAAGCTGTCCAGCTGC	437		
QY	61 CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrThrValLeu	80		
DB	438 TGTTCCTCCCTTCAGGACATATTTTGGCATGCTGTTCACAGATGATACCATCTCTCTA	497		
QY	81 TrpAsnThrGluAsnGlyGlnMetLeuAlaValMetGluGlnProSerGlySerProVal	100		
DB	498 TCGAATACGAAATGCAATGATGCTGGCGACGATGGAACGACCTTAGTGGCAGCCCTTG	557		

QY	101	ArgVal	CysGlnPheSerProAspSerThrCysLeuAlaSerGlyValAlaAspGlyThr	120
Db	558	AGGGTTGGCCAGT	TTTCCCGACTCCACGGTGTGGCATCAGGGGACGCTGATGAGACT	617
QY	121	ValValLeu	TPAaaAlaGlnSerTyrLysLeuTyrArgCysGlySerValLysAspGly	140
Db	618	GTGGTTTGTGGAA	TGCACACTCTATCAAAATTAATATGATGTGTGTAATGTTAAAGATGCG	677
QY	141	SerLeuAla	CysValaPheSerProAsnGlySerPheValThrGlySerSerCys	160
Db	678	TCTTGGCGGCGAT	GTGATTTTCTCTTAATGGAAGCTTCTTGTGCACGTGCTCTCCATGT	737
QY	161	GlyAspLeu	ThrValTPAAspAspLysMetArgCysLeuHisSerGlyLysValaHisAsp	180
Db	738	GGTAGTTTAAACAG	TGGGATGATAAATAGAGTGTCTGCATAGTGAAGAAAGCACAATGAT	797
QY	181	LeuGlyLeu	ThrCysCysAspPheSerSerGlnProValSerAspGlyGlnGlnGlyLeu	200
Db	798	CTTGGAAATTA	CTCTCTCGATTTTCTTTCACGCACTTCTGATGGAAGAAACAGGCTT	857
QY	201	GlnPhePhe	ArgLeuAlaSerCysGlyGlnAspCysGlnValLysIleTrrPileValSer	220
Db	858	CAGTTTTTTCAG	CTGGCATCATGTGTGTCAGAGTTGCCAAGTCAAAATTTGGATTGTTCT	917
QY	221	PheThrHis	IleLeu-----	225
Db	918	TTTACCATTA	CTTAAGCTTTGAAATTAATAATATAAAAGTACACGAGTGGGCACTGTCT	977
QY	225	-----	-----	225
Db	978	CCTGTTCGCTTG	CTGCTTTTCCCATGATGGCAGATGCTAGTCAAGGCTCAGTGAT	1037
QY	225	-----	-----	225
Db	1038	AAGTCTGTCA	TATATATGATTAATACTAATACTGAGATATACTTCAACATTGACTCAGCAC	1097
QY	225	-----	-----	225
Db	1098	ACCAGGTA	TATGCACACTTGTGCTTTGGACCTTAATACCTTTTACTTGCATCGTTCA	1157
QY	226	-----	-----AlaArgArg	228
Db	1158	ATGCACAAA	CAGTGAACATCTGGCAATTTGACTGCGAAACACTTTCGCAAGAGAGC	1217
QY	229	ThrGluHis	GlnLeuLysGlnPheThrGluAspTrrSerGluGluValValSerThrTrr	248
Db	1218	ACAGAACTCA	CGCTGAAACAAATTAACGAAATTTGTGACAGAGGAGATGTCTCAACATGG	1277
QY	249	LeuCysAla	GlnAspLeuLysAspLeuValGlyIlePheLysMetCAsnAsnIleAspGly	268
Db	1278	CTTTGTGCACA	AGATTTTAAAGATCTTGTGTGTAATTTTCAAGATGAATTAACATTTGATGGA	1337
QY	269	LysGlnLeu	LeuAsnLeuThrLysGlnSerLeuAlaAspAspLeuValGlnSerLeu	288
Db	1338	AAAGAACTGT	GAAATCTTAACAAAGAAAGTGTGCTGATGATTTGAAAAATTTGAATCTCTA	1397
QY	289	GlyLeuArg	SerLysValLeuArgLysIleGlnGlnLeuArgTrrLysValLysSerLeu	308
Db	1398	GGACTCGCTAG	TAAGAGTCTAGGAATTAAGAGCTCAGGACCAAGGTTAAATCCCTT	1457
QY	309	SerSerGly	IleProAspGlnPheIleCysProIleThrArgGlnLeuMetLysAspPro	328
Db	1458	TCTTCAGGAAT	TTCCTGATTAATATATATGTCATTACTAGAGAATTAATGAAGATCCG	1517
QY	339	ValIleLeu	AspAspGlyTyrSerTyrGlnLysGlnAlaMetGluAsnTrrIleSerLys	348
Db	1518	GTCAATCGCAT	AGATGCTTATCATATGAAAGAGCAATGGAATAATTTGATCAGCAAA	1577
QY	349	LysLysArg	TrrSerProMetThrAsnLeuValLeuProSerAlaValIleLeuThrProAsn	368
Db	1578	AAGAAACGTACA	AGTCCATGACAAATCTTGTCTTCTTCAGCGGATCTTACACCAAT	1637

QY	369	ArgThrLeuLysMetValAlaAspArgTrpLeuGluThrHisSGlnLys	384
Db	1638	AGGACTGTGAAAATGGCCATCATATGATGCTGGAGACACACAAAAG	1685
RESULT 8			
LOCUS	CQ729621	1254 bp	DNA
DEFINITION	Sequence 15555 from Patent WO02068579.		
ACCESSION	CQ729621		
VERSION	CQ729621.1	GI:42301208	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
REFERENCE			
AUTHORS	Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.		
TITLE	Kits, such as nucleic acid arrays, comprising a majority of		
	humanexons or transcripts, for detecting expression and other uses		
	thereof		
JOURNAL	Patent: WO 02068579-A 15555 06-SEP-2002;		
FEATURES			
source	1..1254		
	Location/Qualifiers		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Alignment Scores:			
Pred. No.:	6,13e-158	Length:	1254
Score:	1823.50	Matches:	357
Percent Similarity:	85.95%	Conservative:	4
Best local Similarity:	85.00%	Mismatches:	20
Query Match:	89.08%	Indels:	39
DB:	6	Gaps:	4
US-10-077-111-13 (1-384) x CQ729621 (1-1254)			
QY	1	MetValLysLeuLeuLeuHisThrLeuAlaAspHisGlyAspAspValAsnCysCysAlaPhe	20
Db	1	ATGTGGAAACGATTCACACATTAAGCTATCATGTGACAGATGCAATGCTGCTGCTTC	60
QY	21	SerPheSerLeuLeuAlaThrCysSerLeuAspLysThrIleArgLeuTySerLeuArg	40
Db	61	TCTTTTCCCTGGCTGCTACTGCTCTTGGACAAACAAATCGCCTGACTGTTAGCT	120
QY	41	AspPheThrGluLeuProHisSerProLeuLysPheHisThrTyAlaValHisCysCys	60
Db	121	GACTTACTGAACTGCCAATCTTCATGAAAGTTTCACTTACTGATGCTGCACGCTGC	180
QY	61	CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrThrValLeu	80
Db	181	TGTTTCTCCCTTCAGGACATATTTTGGCATCTGTTTCAACAGATGTAACACTGTCTTA	240
QY	81	TrpAsnThrGluAsnGlyGlnMetLeuAlaValMetGluGlnProSerGlySerProVal	100
Db	241	TGGATATCTGAAAATGGAACAATGCTGGCAGATGAGAAACAGCCTAGTGGCAGCCCTGTG	300
QY	101	ArgValCysGlnPheSerProAspSerThrCysGluAlaSerGlyAlaAlaAspGlyThr	120
Db	301	AGGGTTTCCACGTTTTTCCCAAGACTCCACGTTTGGCATCAGGGGCGCTGATGAACT	360
QY	121	ValValLeuTrpAsnAlaGlnSerTyTrpLysLeuTyArgCysGlySerValLysAspGly	140
Db	361	GTGGTTTGTGGATGTCACAGTCATACAAATATATTAAGTGTGATGATTTAAAGATGCC	420
QY	141	SerLeuAlaAlaCysAlaPheSerProAsnGlySerPhePheValThrGlySerSerCys	160
Db	421	TCCTTGGGGGCGATGTGCATTTTCTCTAATGGAACCTTCTTGTGCATCGGCTCCATGT	480
QY	161	GlyAspLeuThrValTrpAspAspLysMetArgCysLeuHisSerGluLysAlaHisAsp	180

Accession	Version	Keywords	Source	Organism	Reference	Authors
Db	481	GGTATTTAAACAGTGTGGATGATAAATATGAGGCTGTGCAATAGTGAAGCAACATGAT	540			
Qy	181	LeuGLYL1ethrCysCySaephSesSerGlnProValSer-----	194			
Db	541	CTTGAATATTCACGCTGCTGCAATTTTCTTCAACAGCAGATTTCTGGTTTGAATTAATAT	600			
Qy	195	-----	AspGly 196			
Db	601	AAAAGTACACTGAGTGGGCACTGTGCTCTGCTTGTGCTTTTCCCATGATGGG	660			
Qy	197	GlunGlnGluGlnPhePhe-----	ArgLeu 205			
Db	661	CAGATGTCAGTCTCAAGGATATGTCAACAATTTGCTTTTGCACTTAATACCTTTACTT	720			
Qy	206	AlaSerCysGlyGlnAepCysGlnValLysIleThrIleValSerPhe---ThrHisIle	224			
Db	721	GCTCTCGTTCATATGACCAAAACAGTGAACCTGGCAATTTGACCTCTGGAACCACTTTC	780			
Qy	225	LeuAlaArgAargThrGluHisGlnLeuLysGlnPheThrGluAAPTTPSerGluGluVal	244			
Db	781	CAACCAAGAGAGACACAGAAACATCAGCTCAAGCAATTTACCGAAGATGTGTCAAGAGAGAT	840			
Qy	245	ValSerThrTrpLeuCysAlaGlnAAspLeuLysAspLeuValGlyIlePheLysMetAsn	264			
Db	841	GTCCTCAACATGCTTTGTGTGACACAGATTTAAAGATCTTGTTGGTATTTTCAATGAT	900			
Qy	265	AsnIleAspGlyLysGluLeuLeuAsnLeuThrLysGluSerLeuAlaAspAspLeuLys	284			
Db	901	AACATTTGATGGAAGAAAGACTGTTGATCTTACCAAAAGAAAGTGTGGCTGATGATTTGAA	960			
Qy	285	IleGluSerLeuGlyLeuAArgSerLysValLeuArgLysIleGluGluLeuAArgThrLys	304			
Db	961	ATT-----GGACTGCGTGAAGAGTGTGGGAAATGTGAAGAGCTCAGGACCAAG	1011			
Qy	305	ValLysSerLeuSerSerSerGlyIleProAspGluPheIleCysProIleThrArgGluLeu	324			
Db	1012	GTTAAATCCCTTTCTTCAAGAAATCTCGTAGAATTTATATGTCCAAATPACTAGGAACCTT	1071			
Qy	325	MetLysAspProValIleIleAspAspGlyTyrSerTyrGluLysGluIleMetGluAsn	344			
Db	1072	ATGAAGATCCGCTCATCGCATCGATCGATGCGCTATTATATGATAAAGAAAGCAATGGAAT	1131			
Qy	345	TrpIleSerLysLysLysValArgThrSerProMetThrAsnLeuValIleuProSerAlaVal	364			
Db	1132	TGGATTCAGCAAAAAGAAACGTACCAAGTCCCATGTCAAAATCTTGTTCTTCTTCACGGTA	1191			
Qy	365	LeuThrProAsnAArgThrLeuLysMetAlaIleAsnArgTrpLeuGluThrHisGluLys	384			
Db	1192	CTTACACCAATATGAGACTGTGAATAATGCCCATCATATGATGTGCTGGAGACACACCAAAAG	1251			
RESULT 9						
LOCUS	BC061948	2130 bp	mRNA	linear	ROD 19-NOV-2003	
DEFINITION	Mus musculus RIKEN cDNA 2610014F08 gene, mRNA (CDNA clone MGC:65547.7)					
ACCESSION	BC061948					
VERSION	BC061948.1	GI:3803926				
KEYWORDS	MGC.					
SOURCE	Mus musculus (house mouse)					
ORGANISM	Mus musculus					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Murinae; Mus.					
AUTHORS	1 (bases 1 to 2130) Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuller, G.D., Altschul, S.F., Zeeberg, B., Bluet, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heien, P., Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stepien, T.M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheer, T.E., Brownstein, M.J., Udell, T.B., Yoshizuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullan, S.J., Boeck, S.A., McEwan, P.J., McEwan, K.T., Malek, J.A., Gunaratne, P.H., Richards, S.,					

Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbe, R.A., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E., Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
2 (bases 1 to 2130)
Strausberg, R.
Direct Submission
Submitted (10-NOV-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Susan L. Sullivan, PhD.
CDNA Library Preparation: Reggen, Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nsl.nih.gov
Ahter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-I., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stankiv, P., Thomas, P.J., Touchman, J.W., Tsaur, J., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAC Plate: 127 Row: 1 Column: 1
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.
FEATURES
Source
Location/Qualifiers
1..2130
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="MGC:65547 IMAGE:6312489"
/tissue_type="Olfactory epithelium, neonatal mouse, C57Bl/6"
/clone_1ib="NIH_MGC_129"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1..2130
/gene="2610014F08R1k"
/note="synonym: 1700048E19R1k"
/db_xref="LocusID:72137"
/db_xref="MGI:1919387"
208..1632
/gene="2610014F08R1k"
/codon_start=1
/product="2610014F08R1k protein"
/protein_id="AAH61948.1"
/db_xref="GI:38303927"
/db_xref="LocusID:72137"
/db_xref="MGI:1919387"
/translation="MTRLITLADHDGDDVCCAFSAALLATCSLIDKTRIRYSIDPAE LPYSLKFTHYAAHCCCFSPSGHVLASCSDGTIVLWSSHSHTLVLEPQSGSPRV CFPSPSAVILASGADDSIALMNAQTKLVRCGSYVDSILVACAFSPDGLPVTGSSG

GNLTWDBMRCLHSBKADLCTCCSRSSQPLSGEGIQSYQLASCCDCEIKMAY
TTTRVLFEPFKKSTLSGCAPVLAASFHDCKMLASGVDSVLIHIGIPQSVLHTL
TQTRVITTCAPAPNTLLATGSMKXTVIMQFDLETPOGSMNDPLKFTBESBED
VAVWIAQGLDPLVGI FRANNIDGKELHLTRESLAGDKISLGRSVLSIELR
AKMDSSGIPDEPFCPTRELMKPOVLASDGYSYREAMEBIMHKKRTSPMTNAL
BLVLTNPRTLMCAINRWLETHEK"
217..1131
/gene="2610014F08R1k"
/note="WD40; Region: WD40 domain, found in a number of eukaryotic proteins that cover a wide variety of functions including adaptor/regulatory modules in signal transduction, pre-mRNA processing and cytoskeleton assembly"
/db_xref="CCD:cd00200"
1192..1383
/gene="2610014F08R1k"
/note="SAM; Region: Sterile alpha motif."
/db_xref="CCD:cd00166"
1408..1629
/gene="2610014F08R1k"
/note="U-box; Region: U-box domain. This domain is related to the Ring finger pfam00097 but lacks the zinc binding residues"
/db_xref="CCD:pfam04564"
ORIGIN
Alignment Scores:
Pred. No.: 1,236-138 Length: 2130
Score: 1617.00 Matches: 312
Percent Similarity: 72.21% Conservative: 31
Best Local Similarity: 65.68% Mismatches: 40
Query Match: 78.9% Indels: 92
DB: 10 Gaps: 3
US-10-077-111-13 (1-384) x BC061948 (1-2130)
QY 1 MetValLysLeuLLeIshThrLeuAlaAaPhIsgLYAaPbAyaLaAsnCysCyAlaPhe 20
DB 208 ATGTGAGGCTGATTCACACGCTGCGTGAATCAGGAGATGAGCTGCTGCGCTTC 267
QY 21 SerPheSerLeuLeuAlaThrCysSerLeuAaPbAyaLThrILaArgLeuYrSerLeuAaG 40
DB 268 TCGGCTGCGCTCTGCGCACCTGCTCTTGACAAAGACATCGGCTGATCTCCCTAAGT 327
QY 41 AaPheThrGluLeuProHisSerProLeuLysPheIshThrTyraLaValHisCysCys 60
DB 328 GACTTTCGTAAGTCCGCTGATCTCCCGTGAAGTTCCACACCTATCTCTCACTGCTGC 387
QY 61 CysPheSerProSerGlyHisIshLeuAlaAaSerCysSerThrAaPglYThrValLeu 80
DB 388 TGTTCCTCACCTCAGACACGTTTAAAGATCGTCTCGACAGACGAGACCGGTGCTG 447
QY 81 TrpAaThrGluAaAaGlyGlnMetLeuAlaValMetGluGlnProSerGlySerProVal 100
DB 448 TGGAGCTCCACACGCGACACACCTGACCGGTGTGAGACACGCGGTGACGCTGTCG 507
QY 101 ArgValCysGlnPheSerProAaPbSerThrCysLeuAlaAaSerGlyValAlaAaPglYThr 120
DB 508 CGCGTGTGCTGCTTCCCAAGACTGCTGCTACTGCTGAGGAGGCTGCCATGATGATCC 567
QY 121 ValValLeuTrpAaAaIaGlnSerTyryLysLeuTyraGlySerValIyAaPbGly 140
DB 568 ATGCTTGTGGATGACACGACATATCAAACTATATAGATGTGTGTCAAGATGATAC 627
QY 141 SerLeuAlaAaCysAlaAaPbSerProAaAaGlySerPheValThrGlySerSerCys 160
DB 628 TCATTGTGTGCGCTGCTTCCCAAGAGGCTGCTTGTCTGCTGCTGCTGCTGCTGCTG 687
QY 161 GlyAaPbLeuThrValItrpAaAaPbLysMetAaGlyCysLeuHisSerGlyLysValAaHIsaAaP 180
DB 688 GGGAGACTTGACACTGTGGATGACAGATGAGGTGTCTTACACGACAAAGGCGCCAGAT 747
QY 181 LeuGlyIleThrCysAaPbPheSerSerGlnProValSerAaPglYglnGlnGlyLeu 200

```

Db      748 CTCGGGATCACCTCTGAGCGCTTTCCACAGCCTCTCTGGCGGAGAA---GGCCTC 804
Qy      201 GlnhphearglaualaserCygluInapCygluInapValysIleThrIleValaser 220
      805 CAGCTTACAGGTTGGCGCTCATGTGGTCAAGACTGTGAATCAACTCTGGGCTGTACT 864
Qy      221 PheThr-----Direct Submision 222
      865 ATTACCGCGTCTTAGGCTTTGAATTAATAATAAAGACACTAAGTGGCAGTGGCC 924
Qy      222 ----- 222
Db      925 CCTGTCTGCGCTGTGCTTTTTCACATGATGAAAGATCTTGCACTGGGCTCAGTGAT 984
Qy      223 -----HisIleleuAlaArgArg 228
      985 AAATGTGTCATCATATACATGATGTGGCCCTGAGAGTGTCTACACACGCTGACTCAGCAT 1044
Qy      229 ThrGluHis----- 231
      1045 ACCAGGATGTGTAAGACTTGTGCGCTTGCACCCAACTCTCTTACTTGTCTACTGTTCA 1104
Qy      231 ----- 231
Db      1105 ATGACAAAGACAGTGAACATTTGGCAGTTTGACCTGAAACACCTTGCCCAAGAGATG 1164
Qy      232 -----GlnleuLyrgInpHeThrGluAspTrpSerGluGluValaserThrTrpLeu 249
      1165 AACGACCGGCTGAACATTTCACTGAAAGAAAGTGTGCTGATTTGAAATGATCTCTAGG 1224
Qy      250 CysIleaglInapLeuLyAspLeuValGlyIlePheLyMetAsnAsnIleAspGlyLy 269
      1225 CGTGCTCAGGCTTGAAGACCTGCTCGTATTTTCAGGGCAACACATGATGAGGAA 1284
Db      270 GlnleuLeuAsnleuThrIleGluSerLeuAlaAspAspLeuLyIleGluSerLeuGly 289
      1285 GAATCATTTGATCTCAACAAAGAAAGTGTGCTGATTTGAAATGATCTCTAGG 1344
Qy      290 LeuArgSerLyValleuArgLyIleGluGluLeuArgThrLyValLySerLeuSer 309
      1345 CTGGCGACGCAAGCTCTGAGGATTTGMAAGCTCAGGGCAAGATGATGATCTCTCT 1404
Qy      310 SerGlyIleProAspGluPheIleCysProIleThrArgGluLeuMetLyAspProVal 329
      1405 TCCGGATCCCTTGAAGAGTATCTGCCCATAACAGAGAACTCATGAAGACCCGCTC 1464
Qy      330 IleIleAspArgGlyLyTrpSerLyGluLyGluIleMetGluAsnTrpIleSerLyLy 349
      1465 ATCCATCAGATGGCTACTCTCAGAGAGAGAGAAAGAAAGTGTGATCACAAGAG 1524
Db      350 LyAsnArgThrSerProMetThrAsnLeuValleuProSerAlaValleuThrProAsnArg 369
      1525 AAGCGTACGAGCCCCCATGACAAATTTGGCTCTCCCTTCACTGATCTGACCCCAACAG 1584
Qy      370 ThrleuLyMetAlaIleAsnArgTrpLeuGluThrHisGluLy 384
      1585 ACACCTGAAGATGGCCATCAACCGATGGCTGAGAGCAGCAGAGAG 1629

```

```

RESULT 10
AJ719461      3818 bp      mRNA      linear      VRT 30-SEP-2004
LOCUS      AJ719461
DEFINITION      Gallus gallus mRNA for hypothetical protein, clone 2121.
ACCESSION      AJ719461
VERSION      AJ719461.1 GI:53127473
KEYWORDS      ORF1.
SOURCE      Gallus gallus (chicken)
ORGANISM      Gallus gallus
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE
1      Caldwell, R.B., Kierzek, A.M., Arakawa, H., Bezubov, Y., Zaim, J.,
AUTHORS

```

```

TITLE      Fiedler, P., Kuter, S., Blagoderaki, A., Kostovska, D., Koter, M.,
Plachy, J., Carnici, P., Hayashizaki, Y. and Buerstedde, J.M.
JOURNAL      Published
REFERENCE      2 (bases 1 to 3818)
AUTHORS      Caldwell, R.B.
TITLE      Direct Submission
JOURNAL      Submitted (20-MAY-2004) Caldwell R.B., GSF - Forschungszentrum,
Instituc fuer Molekulare Strahlenbiologie, Ingolstaedter Landstr.
1, D-85764 Neuherberg, GERMANY
FEATURES
source
1..3818
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="CB"
/db_xref="taxon:9031"
/clone="2121"
/cell_type="lymphocyte"
/tissue_type="bursa"
/clone_lib="rikeni"
/dev_stage="2 weeks old"
1..3818
/locus_tag="RCJMB04_2121"
<1..341
/locus_tag="RCJMB04_2121"
342..1469
/locus_tag="RCJMB04_2121"
/notes="ORF1"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAQ31120.1"
/db_xref="GI:53127474"
/translation="MATLIRTLTDSDDVNCAPSSCLATGSLDITIRIYSLDPTF
LPYSLKAGTVAHCCCFSPSGHTLASCTDGTATITWDSDEMLAVLEQPTSPVAV
CRSPSPSTVLSGADGSLVLMVHSMKPYRSGNVDGSLVACAFSGGCFPTGSSC
GDLTWMDDKRCCKNEKXHDGLVTCDSISHPVSDGSHAGCFOMASCGODNKKWF
ILPADPLGELRYKCTLSGHSRAVLTCAPVSDYOMVSGSDVCVITTEYNTNIIHT
LSQHTRIYTTTCARAPCSLPLATSGSMKTHIWKIDNKQPCAGNTIENDSKIRIDENMS
EDVSAWLCAGQFAELVGLFRANNIDGKELVNLRLSLHE"
1470..>3818
/locus_tag="RCJMB04_2121"
ORIGIN
Alignment Scores:
Pred. No.:      3,18e-124      Length:      3818
Score:      1464.00      Matches:      281
Percent Similarity:      67.30%      Conservative:      40
Best Local Similarity:      58.91%      Mismatches:      62
Query Match:      71.52%      Indels:      94
DB:      5      Gaps:      3
US-10-077-111-13 (1-384) x AJ719461 (1-3818)
Qy      1 MetValLyLeuIleHisThrLeuAlaAspHisGlyAspAspValAsnCySValaPhe 20
      342 ATGGCGACATTAATTCACACTTTAAAGATCAAGTATATGCAACTCTGCGCTTC 401
Db      21 SerPheSerLeuLeuAlaThrCysSerLeuAspLyThrIleArgLeuTrpSerLeuArg 40
      402 TCATCATCGTGTGGTGTACTTGTCTTGGACAAACAATTTGGTATTTACTCTTAAAT 461
Qy      41 AspPheThrGluLeuProHisSerProLeuLyPheHisThrTrpAlaValHisCyS 60
      462 GACTTCACAGCTCCCATCTACTCACCTTGAAGAGTCATCATATGCTGTCCACTGTGC 521
Db      61 CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrTrpValleu 80
      522 TGCTTCTTCTCATCAGACCACTTATGCTTCGTTGTTCAACACAGCGGTCTTACCATAT 581
Qy      81 TrpAsnThrGluAsnGlyGlnMetLeuAlaValMetGluGlnProSerGlySerProVal 100
      582 TGGGACACTAGCGATGTGTGGATGCTGGCAGTGTGAGCAGCCACTGGCAGTCTGCTC 641

```



```

Qy 101 ArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyAlaAlaAspGlyThr 120
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 642 AGAGTGTGCGGCTTTTCTCCGAGTCCACTTACTGCTGCAGGTGACGTGATGAGT 701

Qy 121 ValValLeuThrPheAlaGlnSerTyrLysLeuTyrArgCysGlySerValLysAspGly 140
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 702 GTGCTTCTTGGAAACGTGACCTCAATGAAATTTCTACAGATCTGGGAAATGTTAAAGATGT 761

Qy 141 SerLeuAlaAlaCysAlaPheSerProAsnGlySerPhePheValThrGlySerSerCys 160
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 762 TCTTTGGTGGCTGCTGCTTTCTCCAGGTGAAATTTCTTTGTTACTGATCTCATGT 821

Qy 161 GlyAspLeuThrValTTPAspAspLysMetArgCysLeuHisSerGlyValAlaHisAsp 180
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 822 GGTGATCTGACTGTTTGGATGATTAATGATGATGCTCTGTATATGAAAAAGCATGAT 881

Qy 181 LeuGlyLysThrCysCysAspPheSerSerGlnProValSerAspGlyGlnGlyLeu 200
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 882 CTGGGCTTACCTGCTGTGATATTCTTCACTCCACTTCTGATGGGAAACATGCAATCC 941

Qy 201 GlnPhePheArgLeuAlaSerCysGlyGlnAspCysGlnValLysIleTPIleValSer 220
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 942 GGATGCTTCCAGATGAGCATCTTGTGTCAAGATATATAGATCAAAAGTCTGTTATTGTG 1001

Qy 221 PheThrHisIleLeuAlaArgArgThrGlu----- 230
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 1002 TTTCGACAGCTTTAGGTGGTGAATGAGATATATATATGACATTGAGTGGACATTCTGCC 1061

Qy 230 ----- 230
Db 1062 CCAGTTCTGACTGTGCAATTTCTTATATGAGGAGATGTTGTTTCAGGGCTGTGGAC 1121

Qy 231 -----HisGlnLeuLysGlnPhe 236
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 1122 AAGTGTCTATATATATGAGACTAATATCTGGCAATATCTCTTCACTTGTCTCAGCAT 1181

Qy 237 Thr----- 237
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 1182 ACTAGATATGTCACAACTTGCTGCTCATGTAGTCTCTTCTTGCACGGGTTCA 1241

Qy 237 ----- 237
Db 1242 ATGATATAAACTGTGACATATGAGAACTGACAACTCTGTGACGAAACACT 1301

Qy 238 -----GluAspTrpSerGluGluValLysThr 247
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 1302 ATGAAAAATGACTTAAGATTAGAACTGCTGAGAACTGCTCAGAGATGATGTTTCAGCC 1361

Qy 248 TrpLeuCysAlaGlnAspLeuLysValGlyLysPheLysMetAsnAsnLysAsp 267
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 1362 TGGCTTTGTGCAAGGCTTGCAGAACTTGTGGCTTTTCAAAAGCAATATCATGAT 1421

Qy 268 GlyLysGlnLeuLeuAsnLeuThrLysGlnSerLeuAlaAspAspLeuLysIleGluSer 287
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 1422 GGCAGAAAGAACTCGGAATCTTACAAAGAAAGCTGATTCATGATATAAAATG---AGT 1478

Qy 288 LeuGlyLeuArgSerLysValLysArgLysIleGlnGluLeuArgThrLysValLysSer 307
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 1479 CTCTGGCTGGCAATAATCTCTCAAGAAATGCAAGAGCTGAGAGATGAAATGCTCTCT 1538

Qy 308 LeuSerSerGlyLysProAspGluPheLysCysProLysThrArgGluLeuMetLysAsp 327
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 1539 GTTCTGCTGTGCTGCTGATGATTTCTTATGCTTAAACAGGAGGCTCATTAAGAT 1598

Qy 328 ProValLysLeuAspAspGlyLysSerTyrGlyLysGluAlaMetGluAsnTrpLysSer 347
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 1599 CTTGTCAATTCAGACAGATGGGTATTTCTATGAAAGAGGCAATGAGAAATTTGATCAGC 1658

Qy 348 LysLysLysArgThrSerProMetThrAsnLeuValLeuProSerAlaValLysLeuThrPro 367
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 1659 AATTAACAGAGATTAATGCTCCATGACAAATCTTCTCTCCAGTCTTGTGCTTACACC 1718

```

```

Qy 368 AenArgThrLeuLysMetAlaIleAenArgTrpLeuGlnThrHisGlnLys 384
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 1719 AACAGGACACTGAAATGCTATCAGCCCTGCTGAGACTCAGCAGAAAA 1769

RESULT 11
BC050792
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
1 (bases 1 to 1547)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Uedtin,T.B., Tomshiyuki,S.,
Carninci,P., Prange,C., Rana,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullaby,S.J., Bosak,S.A., McEwan,P.J.,
Mckernan,K.J., Malek,J.A., Gunaratne,P.H., Richardson,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hultk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shewchenko,Y.,
Bouffard,G.G., Blakeley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Sklarski,U., Smallus,D.B.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
223888257
12477932
2 (bases 1 to 1547)
Strausberg,R.
Direct Submission
Submitted (09-APR-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. Jonathan Kuo, NIMH
cDNA Library Preparation: Michael Brownstein / Ted Uedlin
laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@cgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Boedeltje, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Heiao, Martin Krzywinski, Keta Kutche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice Mcleavy, Steven
Ness, Pawan Parbhu, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smallus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranda Tasi, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.
clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
Series: IRAL Plate: 46 Row: a Column: 6
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis.
Location/Qualifiers

```


source

1. .1547
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6742025"
/tissue_type="Testicle, mouse"
/clone_id="N1H_MGC_169"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"

gene

<1. .1547
/gene="2610014F08Rik"
/note="synonym: 1700048E19Rik"
/db_xref="LocusID:72137"
/db_xref="MGI:1919387"

<1. .1442
/gene="2610014F08Rik"
/codon_start=3
/product="2610014F08Rik protein"
/protein_id="AAH50792.1"
/db_xref="GI:30048137"
/db_xref="LocusID:72137"
/db_xref="MGI:1919387"
/translation="PSSGSPSPSPGQCRGLDLTRRCVTSSTEPRRGGAARRDSGS
GCBSCCLCVNMRVLIHTLADHDVSCAFSALLATSLDITLISLSPAEIP
YSPLEHTYAVHCCFSPSGHVLASCTDGTVLMSHSGHTLVLEPGSGPRVCC
FSPSAYLAGSADGSIALNMAOTYKLYRGSVSDSLVCAIPSDGLFVTSGGD
LYWDPMRCLISERKADLITCCSPSSQPSGSGEIQSVQASCGGDCIKIMAYTI
TRYGELETKSTLSHGCAVPYACAFSHDGMKLSGSDVSVIIGHIGPVSHTLTQ
HTRVTTCAFPNTLLATGSMKTANIWQFDLETPOGSMNDPLKHTFMESEEDVS
VMLRAQGLIEDLVGIFRANNIDGKELHLHTKESLAGDILKISLGRSKVLSIEELRAK
MDSISGIPDEFICPTIRELMKDPVIASG"

CDS

ORIGIN

Alignment Scores:

Pred. No.: 2.25e-116 Length: 1547

Score: 1373.00 Matches: 272

Percent Similarity: 67.18% Conservative: 33

Best Local Similarity: 59.91% Mismatches: 45

Query Match: 67.07% Indels: 104

DB: 10 Gaps: 4

US-10-077-111-13 (1-384) x BC050792 (1-1547)

Qy 1 MetValysLeuIleHsThrlLeuAlaAspHisGlyAspAspValaAncysCysAlaPhe 20

Db 171 ATGATGAGGCTGATTCACAGCTGCTGATCAGCTGATGACGCTGCTGCTGCTTC 230

Qy 21 SerPheSerLeuLeuAlaThrCysSerLeuAspThrIleArgLeuYrSerLeuArg 40

Db 231 TCGGCTGCTCTCCCTGGCACCCTGCTCTTGACAAAGACCATCCGCTGCTCTCACT 290

Qy 41 AspPheThrGluLeuProHisSerProLeuLeuAspPheHisThrYrAlaValHisCysCys 60

Db 291 GACTTCTGTAAGTCGCGTACTCCCGCTGAAGTTCCACACCTATGCTGCTCACTGCTGC 350

Qy 61 CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrThValLeu 80

Db 351 TGTTCACCTCAGACAGCTTTAGCATGCTGCTGACAGACGAGGACGACGCTGCTG 410

Qy 81 TrpAsnThrGluAsnGlyMetLeuAlaValMetGluInProSerGlySerProVal 100

Db 411 TGAAGCTCGACAGCAGACACACCTTACCTGTGAGACAGCCGGGTGGACGCTCTGTG 470

Qy 101 ArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyAlaAlaAspGlyThr 120

Db 471 CGGCTCTGTTGCTTTCCCGAGACTGCTCACTTACGATGAGGCTGCGCATGAGATCC 530

Qy 121 ValValLeuTrpAsnAlaGlnSerThrYrLeuLeuYrArgCysGlySerValAspGly 140

Db 531 ATTGCTTTGAGATGACAGACATACAACTATATAGTGTGTATGTCMAAGATAGC 590

Qy 141 SerLeuAlaAlaCysAlaPheSerProAsnGlySerPhePheValThrGlySerSerCys 160

Db 591 TCATTGTGCTGCTGCTGCTTTCTCCCATGAGAGGCTCTTTGTCACTGCTCTCGGC 650

Qy 161 GlyAspLeuThrValTrpAspAspLysMetArgCysLeuHisSerGlyLysAlaHisAsp 180

Db 651 GGGGACTTACAGTGGATGACAGATGAGATAGGTGTCTTACACAGCAGAAAGGCGACGAT 710

Qy 181 LeuGlyIleThrCysCysAspPheSerSerGlnProValSerAspGlyGlnGlyLeu 200

Db 711 CTGGGATTCACCTGCTGCTGCTCTTCTCTCAAGCTCTCTCTGCGGAGAA---GGCTC 767

Qy 201 GlnPheAspLeuAlaSerCysGlyGlnAspCysGlnValIleTrpIleValSer 220

Db 768 CAGCTTACAGATTTGGCGCTGATGTGTCAAGACTGTGAATCAAACTGCGCTGTACT 827

Qy 221 PheThr 222

Db 828 ATTACCGTGTCTTAGCTTTGAATTAAATATATAAGACACTAAGTGGCAGCTGGCC 887

Qy 222 222

Db 888 CTTGTTCTGCTGCTGCTTTTTCATGATGAAAGATGCTTGATCGGGGTGATGAGAT 947

Qy 223 223

Db 948 AAATCTGTCAATCAATGATGATCGGCTCAGAGTGTGTACACACGCTGACTCAGCAT 1007

Qy 229 ThrGluHis 231

Db 1008 ACCAGATATGTACGACTGTGCGTTTGACCCCAACATCTTACTTCTGCTACTGTTCA 1067

Qy 231 231

Db 1068 ATGACAAAGACAGTGAACATTTGGCAGTTGACCTGGAACACCTTGCCAGAGACATG 1127

Qy 232 232

Db 1128 AAGACCCGCTGAACATTTCACTGAAGATGTGTGAGAGGATGTCTCGTGTGCTT 1187

Qy 250 CysAlaGlnAspLeuLysAspLeuValGlyIlePheLysMetAsnIleAspGlyLys 269

Db 1188 CGTCTCAAGCTTGGAGAACCTCGTGTATTTTTCAGGCGCAAAACATCGATGGGAAA 1247

Qy 270 GluLeuLeuAsnLeuThrLysGlnSerLeuAlaAspAspLeuValIleGlnSerLeuGly 289

Db 1248 GAACATATGATCTCAACAAAGAAAGCTGCTGCTGATTTGAATCGAATCTTCAAGG 1307

Qy 290 LeuArgSerLysValLeuArgLysIleGlnGluLeuArgThrLysValLysSerLeuSer 309

Db 1308 CTGCGACAGAAAGTCTGAGGAGTATTTGAAGACTCAGGCGCAAGATGATTCCTCTCT 1367

Qy 310 SerGlyIleProAspGluPheIleCysProIleThrArgGluLeuMetLysAspProVal 329

Db 1368 TCCGGAATCCCTGACAGATTCATCTGCCCAATACAGAGAACTCATGAGAGACCCGCTC 1427

Qy 330 IleAlaSerAspGlyLysSerThrGlyLysGluAlaMetGlu 343

Db 1428 ATGCAATCAGGCTTACTAGCAGGTCAAAAGAAATCTCTATATATATGAAA 1487

Qy 344 344

Db 1488 GAATTGTGATCCACAGAAATTAAGAAAGAAAAA 1529

RESULT 12

AX077675 1291 bp DNA linear PAR 22-FEB-200

LOCUS Sequence 130 from Patent WO0105970.

DEFINITION AX077675

ACCESSION AX077675

VERSION 1

KEYWORDS GT-13122057

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Chordata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

TITLE
JOURNAL
COMMENT
Wheeler, J., ~~Wheeler~~, Wyman, D., Ye, W.J. and Zody, M.
Direct Submission
Submitted (16-FEB-1999) Whitehead Institute/MIT Center for Genome
Research, 330 Charles Street, Cambridge, MA 02141, USA
On Mar 1, 1999, this sequence version replaced gi:4309777.
All repeats were identified using RepeatMasker: Smit, A.F.A. &
Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>.

FEATURES
source
1. 147021
Location/Qualifiers

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="hrpx.35 A.1"
/clone_1kb="RPC1-11 human BAC library"
complement(382..683)
/rpt_family="AluSx"
repeat_region
1233..1446
/rpt_family="L1MB3A"
4058..4138
/rpt_family="MER96"
complement(4768..5069)
/rpt_family="AluY"
5268..5687
/note="Single-stranded coverage."
5392..5693
/rpt_family="AluSg"
complement(5856..6314)
/rpt_family="L1MC5"
7499..7713
/rpt_family="MIR"
complement(9050..9111)
/rpt_family="L2"
complement(9768..9867)
/rpt_family="AluIo/FRAM"
9927..10125
/rpt_family="AluSg/x"
complement(11036..11131)
/rpt_family="MIR"
12467..12556
/rpt_family="MIR"
complement(13039..13158)
/rpt_family="MIR"
13216..13368
/rpt_family="MER5B"
complement(13631..13682)
/rpt_family="MIR"
complement(14013..14324)
/rpt_family="AluSg"
complement(14881..15068)
/rpt_family="L2"
complement(15777..15920)
/rpt_family="FLAM_A"
15977..16097
/rpt_family="MER5A"
complement(16111..16422)
/rpt_family="AluSx"
complement(16435..16990)
/rpt_family="L1MB2"
complement(16991..17293)
/rpt_family="AluSc"
complement(17318..17483)
/rpt_family="FAM"
complement(17491..17574)
/rpt_family="L1MB2"
complement(17578..17632)
/rpt_family="MER5B"
18396..18544
/rpt_family="MIR"
complement(18601..18892)
/rpt_family="AluSx"
21078..21102
/rpt_family="(CA)n"

repeat_region complement(21331..21814)
/rpt_family="MLT1D"
repeat_region complement(24151..24326)
/rpt_family="MLT1E1"
repeat_region complement(26673..26977)
/rpt_family="AluSc"
complement(27038..27349)
/rpt_family="AluIb"
27424..27614
/rpt_family="MER3"
28893..29218
/rpt_family="MER7A"
30893..31200
/rpt_family="AluSg"
31226..31363
/rpt_family="MIR"
31425..31608
/rpt_family="MIR"
31694..31715
/rpt_family="AT rich"
complement(32264..32559)
/rpt_family="MER7A"
34027..34056
/rpt_family="(TTTG)n"
complement(34057..34340)
/rpt_family="AluSc"
complement(34350..34643)
/rpt_family="AluSx"
complement(35919..36110)
/rpt_family="MER34B"
36113..36414
/rpt_family="AluSg"
complement(36626..36713)
/rpt_family="MER34B"
37317..37614
/rpt_family="AluSx"
complement(39201..39613)
/rpt_family="L1MB3"
complement(40549..40933)
/rpt_family="MLT1G1"
41531..41566
/rpt_family="(TTTTG)n"
43451..43494
/rpt_family="(TC)n"
complement(43767..44074)
/rpt_family="AluSg"
complement(44084..44352)
/rpt_family="AluIb"
complement(44359..44541)
/rpt_family="L1MA6"
44542..44872
/rpt_family="AluSg"
complement(44873..45183)
/rpt_family="L1MA6"
complement(45708..46376)
/rpt_family="L1MB2"
complement(46377..46677)
/rpt_family="AluSx"
complement(46678..46940)
/rpt_family="L1MB2"
complement(48115..48193)
/rpt_family="MER5B"
complement(48485..48681)
/rpt_family="MIR"
49373..49660
/rpt_family="AluSx"
complement(50181..50391)
/rpt_family="MER58A"
52831..52908
/rpt_family="MIR"
57191..57478
/rpt_family="AluSc"
57518..57795

Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rivers, C., Rodkey, T., Rojars, A., Rose, M., Rose, R., Ruiz, S. J.,
Sanders, W., Savery, G., Scherer, S., Scott, G., Shattman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C. D., Smajls, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorrell, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Swatek, A., Tabot, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
Valas, R., Vera, V., Villaseña, D., Waldron, L., Walker, B., Wang, J.,
Wang, O., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
Weinstock, G. and Gibbs, R. A.

Direct Submission
Unpublished
2 (bases 1 to 245795)
Worley, K. C.

Direct Submission
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 245795)
Rat Genome Sequencing Consortium.

Direct Submission
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

On May 10, 2003 this sequence version replaced gi:24819739.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GFNB
Center clone name: CH230-24D13
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 232948 bases at least Q40
Consensus quality: 234943 bases at least Q30
Consensus quality: 236332 bases at least Q20
Estimated insert size: 240759; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

1 245795: contig of 245795 bp in length.
Location/Qualifiers
1. 245795
/organism="Rattus norvegicus"
/mol_type="genomic DNA"

FEATURES
source

/db xref="taxon:10116"
/clone="CH230-24D13"
1. 1225
misc_feature
/note="wgs_contig"
2242..3454
misc_feature
/note="wgs_contig"
243942..245795
misc_feature
/note="wgs_contig"

ORIGIN

Alignment Scores:
Pred. No.: 2,18e-54 Length: 245795
Score: 731.00 Matches: 164
Percent Similarity: 33.98% Conservative: 10
Best Local Similarity: 32.03% Mismatches: 20
Query Match: 35.71% Indels: 318
DB: 2 Gaps: 1

US-10-077-111-13 (1-384) x AC096419 (1-245795)

Qy 1 MetVallylsleuIleHISTHrleuAlaAspHisGlyAspAspValaencyCy6CyAlaPhe 20
Db 137442 ATGGTGAAGCTGATTCACACACTGGCGGATCACGGTATGACCTGCTCGCGCTTC 137383

Qy 21 SerPheSerleuAlaThrCysSerleuAspIleThrIleArgLeuIYrSerleuArg 40
Db 137382 TCATCCACGCTCTCGGCGACCTGCTCTTGGACAGACCATCCGCTGTACTCCTTAAGT 137323

Qy 41 AspPheThrGluLeuProHisSerProleuIleAspHisThrIYrAlaValHisCysCys 60
Db 137322 GACTTCTGTAACGCGCGACTCCCGTTGAAGTTTCACACGTAATGCGCTCCACTGCTGC 137263

Qy 61 CysPheSerProSerGlyHisIleleuAlaSerCysSerThrAspGlyThrIYrValleu 80
Db 137262 AGCTTCTGCGCGTACAGACAGCTTTGCGCTGCTGCGACAGCGGTACTACGTACTG 137203

Qy 81 TrpAnthrGluSenglyGlyImetleuAlaValMetGluGlnProSerGlySerProVal 100
Db 137202 TGGAGCGCGACAGCGGACACACCTTACGCTGTGAGACAGCGGCGTGAAGCCCTGTG 137143

Qy 101 ArgValCysGlnPheSerProAspSerThrCysleuAlaSerGlyAlaIleAspGlyThr 120
Db 137142 CGGTTTGCTGCTTTTGGCGGACTCCACTGACATCAGGCGCTGCAGATGGGTCT 137083

Qy 121 ValValleuTrpAnthrIleGlnSerIYrIYrleuIYr----- 132
Db 137082 GTTGTGTTTGGAACGACATTCATACAAATTATATGTAACGCGGCTTAACCTGCC 137023

Qy 132 ----- 132
Db 137022 TGGCTGGAAGCCATTAAATTAATTAATTTTCTTCACTGTGTAAATAATAGACATAGT 136963

Qy 132 ----- 132
Db 136962 TGAACCTATTAGTGAAATTAATTAAGTCTTGGAATCGGAAACCTTTAGTTAAGGT 136903

Qy 132 ----- 132
Db 136902 TGAATATCTCATGTGCTCCCTGCTGTGTCGACAGCGGCTGTGCCAGTGGCCAA 136843

Qy 132 ----- 132
Db 136842 GGAAGTGGCCCTTTGTAAGTCTAGAGCTTAAGAAAGGACGATGCTTCAAGCCATGTCCAG 136783

Qy 132 ----- 132
Db 136782 GGGTTAAGCTTAAGCTTAACAGCAGCAGCTTAAGAAAGTGTATATGTCTCCCTTTAGAA 136723

Qy 132 ----- 132
Db 136722 ACGATGCTTAAGCTTAAGAGGGTTTTCACAGAGTGAAGTGTGATTTCCACCTGAGCAG 136663

Qy 132 ----- 132


```
----- Project Information
Center project name: GMBU
Center clone name: CH230-18218
----- Summary Statistics
Assembly program: Phred; version 0.990329
Consensus quality: 246672 bases at least Q40
Consensus quality: 252029 bases at least Q30
Consensus quality: 255877 bases at least Q20
Estimated insert size: 247533; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 18 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 12612: contig of 12612 bp in length
* 12613 12712: gap of unknown length
* 12713 45951: contig of 33239 bp in length
* 45952 46051: gap of unknown length
* 46052 50355: contig of 4304 bp in length
* 50356 50455: gap of unknown length
* 50456 82938: contig of 32483 bp in length
* 82939 83038: gap of unknown length
* 83039 286785: contig of 203747 bp in length
* 286786 286885: gap of unknown length
* 286886 297294: contig of 10409 bp in length
* 297295 297394: gap of unknown length
* 297395 298651: contig of 1257 bp in length
* 298652 298751: gap of unknown length
* 298752 299932: contig of 1181 bp in length
* 299933 300032: gap of unknown length
* 300033 301744: contig of 1712 bp in length
* 301745 301844: gap of unknown length
* 301845 303382: contig of 1538 bp in length
* 303383 303482: gap of unknown length
* 303483 304794: contig of 1312 bp in length
* 304795 304894: gap of unknown length
* 304895 306171: contig of 1277 bp in length
* 306172 306271: gap of unknown length
* 306272 307936: contig of 1665 bp in length
* 307937 308036: gap of unknown length
* 308037 309128: contig of 1092 bp in length
* 309129 309228: gap of unknown length
* 309229 310777: contig of 1549 bp in length
* 310778 310877: gap of unknown length
* 310878 312534: contig of 1657 bp in length
* 312535 312634: gap of unknown length
* 312635 313934: contig of 1300 bp in length
* 313935 314034: gap of unknown length
* 314035 315463: contig of 1429 bp in length.
*
* Location/Qualifiers
*
*   source
*     1..315463
*       /organism="Rattus norvegicus"
*       /mol_type="genomic DNA"
*       /db_xref="taxon:10116"
*       /clone="CH230-18218"
*
*   misc_feature
*     1..1591
*       /note="wgs_contig"
*       11123..12612
*         /note="wgs_contig"
*         29558..31035
*           /note="wgs_contig"
*           46052..47308
*             /note="wgs_contig"
*             50456..52615
*               /note="wgs_contig"
*               59364..60781
*                 misc_feature
```

```
----- Project Information
Center project name: GMBU
Center clone name: CH230-18218
----- Summary Statistics
Assembly program: Phred; version 0.990329
Consensus quality: 246672 bases at least Q40
Consensus quality: 252029 bases at least Q30
Consensus quality: 255877 bases at least Q20
Estimated insert size: 247533; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 18 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 12612: contig of 12612 bp in length
* 12613 12712: gap of unknown length
* 12713 45951: contig of 33239 bp in length
* 45952 46051: gap of unknown length
* 46052 50355: contig of 4304 bp in length
* 50356 50455: gap of unknown length
* 50456 82938: contig of 32483 bp in length
* 82939 83038: gap of unknown length
* 83039 286785: contig of 203747 bp in length
* 286786 286885: gap of unknown length
* 286886 297294: contig of 10409 bp in length
* 297295 297394: gap of unknown length
* 297395 298651: contig of 1257 bp in length
* 298652 298751: gap of unknown length
* 298752 299932: contig of 1181 bp in length
* 299933 300032: gap of unknown length
* 300033 301744: contig of 1712 bp in length
* 301745 301844: gap of unknown length
* 301845 303382: contig of 1538 bp in length
* 303383 303482: gap of unknown length
* 303483 304794: contig of 1312 bp in length
* 304795 304894: gap of unknown length
* 304895 306171: contig of 1277 bp in length
* 306172 306271: gap of unknown length
* 306272 307936: contig of 1665 bp in length
* 307937 308036: gap of unknown length
* 308037 309128: contig of 1092 bp in length
* 309129 309228: gap of unknown length
* 309229 310777: contig of 1549 bp in length
* 310778 310877: gap of unknown length
* 310878 312534: contig of 1657 bp in length
* 312535 312634: gap of unknown length
* 312635 313934: contig of 1300 bp in length
* 313935 314034: gap of unknown length
* 314035 315463: contig of 1429 bp in length.
*
* Location/Qualifiers
*
*   source
*     1..315463
*       /organism="Rattus norvegicus"
*       /mol_type="genomic DNA"
*       /db_xref="taxon:10116"
*       /clone="CH230-18218"
*
*   misc_feature
*     1..1591
*       /note="wgs_contig"
*       11123..12612
*         /note="wgs_contig"
*         29558..31035
*           /note="wgs_contig"
*           46052..47308
*             /note="wgs_contig"
*             50456..52615
*               /note="wgs_contig"
*               59364..60781
*                 misc_feature
```

```
----- Project Information
Center project name: GMBU
Center clone name: CH230-18218
----- Summary Statistics
Assembly program: Phred; version 0.990329
Consensus quality: 246672 bases at least Q40
Consensus quality: 252029 bases at least Q30
Consensus quality: 255877 bases at least Q20
Estimated insert size: 247533; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 18 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 12612: contig of 12612 bp in length
* 12613 12712: gap of unknown length
* 12713 45951: contig of 33239 bp in length
* 45952 46051: gap of unknown length
* 46052 50355: contig of 4304 bp in length
* 50356 50455: gap of unknown length
* 50456 82938: contig of 32483 bp in length
* 82939 83038: gap of unknown length
* 83039 286785: contig of 203747 bp in length
* 286786 286885: gap of unknown length
* 286886 297294: contig of 10409 bp in length
* 297295 297394: gap of unknown length
* 297395 298651: contig of 1257 bp in length
* 298652 298751: gap of unknown length
* 298752 299932: contig of 1181 bp in length
* 299933 300032: gap of unknown length
* 300033 301744: contig of 1712 bp in length
* 301745 301844: gap of unknown length
* 301845 303382: contig of 1538 bp in length
* 303383 303482: gap of unknown length
* 303483 304794: contig of 1312 bp in length
* 304795 304894: gap of unknown length
* 304895 306171: contig of 1277 bp in length
* 306172 306271: gap of unknown length
* 306272 307936: contig of 1665 bp in length
* 307937 308036: gap of unknown length
* 308037 309128: contig of 1092 bp in length
* 309129 309228: gap of unknown length
* 309229 310777: contig of 1549 bp in length
* 310778 310877: gap of unknown length
* 310878 312534: contig of 1657 bp in length
* 312535 312634: gap of unknown length
* 312635 313934: contig of 1300 bp in length
* 313935 314034: gap of unknown length
* 314035 315463: contig of 1429 bp in length.
*
* Location/Qualifiers
*
*   source
*     1..315463
*       /organism="Rattus norvegicus"
*       /mol_type="genomic DNA"
*       /db_xref="taxon:10116"
*       /clone="CH230-18218"
*
*   misc_feature
*     1..1591
*       /note="wgs_contig"
*       11123..12612
*         /note="wgs_contig"
*         29558..31035
*           /note="wgs_contig"
*           46052..47308
*             /note="wgs_contig"
*             50456..52615
*               /note="wgs_contig"
*               59364..60781
*                 misc_feature
```

Db 78089 GGAACAGCGTTGCCAGGACAGCCCTTTCGAACCGTCTTCCAGAGAGCTGCTGCTGC 78148
QY 132 ----- 132
Db 78149 ACATGTTTTTCCATAATGATTTGTTGAGATGAGGACAGGGCTTGTGTAATTGCTG 78208
QY 132 ----- 132
Db 78209 ATTGAATGCAGAAATCGAATGCGTCTTATTCATAAGCATGCTTTTGGTAAGAAA 78268
QY 132 ----- 132
Db 78269 TCATTCTGTAATTTTAATGAGGCTTTGTAGGCTTGCTGCTTTTGGCTTTTAAAG 78328
QY 132 ----- 132
Db 78329 CTTGGGGTTTAAATGCTCATATTGACTCGCTATGAAGTTAAGATACAAGAGTGT 78388
QY 132 ----- 132
Db 78389 AGGAGAGTCTCTGGATCTAGGCTCATAGGTAACAGTGAAGCTTTAAGACTCGCTG 78448
QY 132 ----- 132
Db 78449 TACTCTAAGGGGTGTGACGGGTGACATCTGATGCTTATTGTAGGCAACGGTCACT 78508
QY 132 ----- 132
Db 78509 CTTGTTCTGTCAAGCGCGGTGTTCATTCTAAGGACTGTATAGAGGTCCAAAGCT 78568
QY 133 -----ArgCysGlySerValLysAspGlySerLeu 142
Db 78569 CTGAACCCCGGCTCTGCTGTTCTGTTTCAGGTGTGTGTCAGAGACAGCTCATTTG 78628
QY 143 AlaAlaCysAlaPheSerProAngLysSerPhePheValThrGlySerSerCysGlyAsp 162
Db 78629 GTGGCTGTGCGTTTCTCTGACGAGGCGCTCTTGTCACTGGCTCTCGGAGGGGAC 78688
QY 163 LeuThrValLysAspAspLysMetArgCysLeuHisSerGluLysAlaHisAspLeuGly 182
Db 78689 TTGCAGGTGTGGGATGACAGATAGGTGTCTGCACAGCGAAGGACACGACCTCGGG 78748
QY 183 IleThrCysCysAspPheSerSerGlnProValSer 194
Db 78749 ATCACCTGTGACAGCTTTCTCTCACAGCTCTCTCT 78784

Search completed: February 5, 2005, 13:45:30
Job time : 5876 secs